

AUSTRALIAN JOURNAL

OF PHYSICS

VOLUME 6

QC

1

A85

V.8

1955

SCIENCE

PER

Digitized by the Internet Archive
in 2024

AUSTRALIAN JOURNAL OF PHYSICS

VOLUME 8

MELBOURNE
1955

AUSTRALIAN JOURNAL OF PHYSICS

Published by the Commonwealth Scientific and Industrial Research Organization, in collaboration with the Australian National Research Council and the Institute of Physics (Australian Branch), and under the general direction of the Editorial Board (Dr. N. S. Noble (Chairman), Professor J. S. Anderson, Professor Sir Macfarlane Burnet, Professor L. H. Martin, and Professor J. G. Wood). Volumes 1 to 5 of the Australian Journal of Physics and the Australian Journal of Chemistry issued as the Australian Journal of Scientific Research, Series A : Physical Sciences.

Issued quarterly, 30/- per annum

EDITORIAL ADVISORY COMMITTEE

Chairman and Editor : Dr. N. S. Noble

Members : Dr. G. H. Briggs

Professor L. H. Martin

Professor H. C. Webster

All enquiries and manuscripts should be forwarded to :

The Editor,

Australian Journal of Physics,

Commonwealth Scientific and Industrial Research Organization,

314 Albert Street, East Melbourne, C.2, Victoria.

CONTENTS

NUMBER 1, MARCH 1955

	PAGE
The Distribution of Dislocations in Linear Arrays. By A. K. Head and N. Louat	1
Nets Composed of Parts of Circles for the Approximate Solution of Field Problems. By L. Tasny-Tschiasny	8
Sharpening of Observational Data in Two Dimensions. By E. J. Burr..	30
Correcting for Gaussian Aerial Smoothing. By R. N. Bracewell.. ..	54
A Method of Correcting the Broadening of X-Ray Line Profiles. By R. N. Bracewell	61
A General Equation for the Choice of Glass for Cemented Doublets. By W. H. Steel	68
The Determination of Electron Trajectories in the Presence of Space Charge. By D. L. Hollway	74
The Effect of Multiple Scattering on Electron Energy Loss Distributions. By D. F. Hebbard and P. R. Wilson	90
The Energy Loss Distributions of 1 MeV Electrons in Aluminium Foils. By J. A. McDonell, M. A. Hanson, and P. R. Wilson	98
Microwave and Metre Wave Radiation from the Positive Column of a Gas Discharge. By L. W. Davies and E. Cowcher	108
The Characteristics of Tracks in Nuclear Research Emulsions. By A. J. Herz and G. Davis	129
Observations on the Penetrating Component of Extensive Air Showers. By V. C. Officer and P. J. Eccles	136
Radio Echo Observations of Meteors in the Southern Hemisphere. By A. A. Weiss	148
The Gravitational Compression of an Elastic Sphere. By A. Keane ..	167
Stress Relaxation of Wool Fibres in Water at Strains of 5-20 per cent. By B. J. Rigby	176

Short Communications

A Proposed Modification to the C.S.I.R.O. Mark I Computer. By B. E. Swire	184
Multiple Images in Reflecting Microscopes. By W. H. Steel	186
Cosmic Rays and Air Mass Effects at Macquarie Island. By R. M. Jacklyn	190

NUMBER 2, JUNE 1955.

PAGE

The Sampling Distributions of Statistics Derived from the Multivariate t -Distribution. By E. A. Cornish	193
Chord Construction for Correcting Aerial Smoothing. By R. N. Bracewell	200
The Transport of Heat through Narrow Channels by Liquid Helium II. By P. G. Klemens	206
A Differential Refractometer of High Sensitivity. By P. G. Guest ..	212
The Effects of Collimation and Oblique Incidence in Length Interferometers. I. By C. F. Bruce	224
The Effects of Collimation and Oblique Incidence in Length Interferometers. II. By B. S. Thornton	241
Theory of Tracks in Nuclear Research Emulsions. By J. M. Blatt ..	248
The Influence of Noise on Radar Meteor Observations. By C. D. Ellyett and G. J. Fraser	273
Diffusion Coefficients from the Rate of Decay of Meteor Trails. By A. A. Weiss	279
A Test for Singularities in Sydney Rainfall. By E. J. Hannan ..	289
The Fitting of Truncated Type III Curves to Daily Rainfall Data. By S. C. Das	298

Short Communications

Note on Plasma-Electron Oscillations. By K. G. Emeleus and T. K. Allen	305
A Variable Energy Cyclotron. By D. E. Caro, L. H. Martin, and J. L. Rouse	306
Excitation Studies of Reactions Occurring in the Proton Bombardment of ^{10}B . By J. W. G. Wignall	310
Neutrons Emitted in the Deuteron Bombardment of Carbon. By J. R. Bird	314
Corrigendum	318

NUMBER 3, SEPTEMBER 1955

	PAGE
Hydromagnetic Stability of a Current Layer. By R. E. Loughhead ..	319
Correcting for Running Means by Successive Substitutions. By R. N. Bracewell	329
The High Temperature Excitation of Helium. By J. T. Jefferies ..	335
Galactic Survey at 400 Mc/s between Declinations -17° and -49° . By R. X. McGee, O. B. Slee, and G. J. Stanley	347
The Observation and Interpretation of Radio Emission from some Bright Galaxies. By B. Y. Mills	368
The Radar Determination of Meteor Showers in the Southern Hemisphere. By C. D. Ellyett and K. W. Roth	390
The Distribution of the Ratio of Two Quadratic Forms. By G. S. Watson	402

Short Communications

The Design of a Multiple-wire Proportional Counter and its Use in the Study of the α -Particles from the Reaction ${}^7\text{Li}+p$. By A. C. Riviere and P. B. Treacy	408
On Starting Routines for the C.S.I.R.O. Mark I Computer. By G. W. Hill and T. Pearcey	412
Eigen Oscillations of Compressible, Ionized Fluids. By R. E. Loughhead	416
A Proportional Counter with Grid Control. By R. J. Norman ..	419

NUMBER 4, DECEMBER 1955

PAGE

The Solution of Linear Simultaneous Equations by Matrix Iteration. By J. Guest	425
The Statistical Design of an Experiment to Test the Stimulation of Rain. By P. A. P. Moran	440
Photodisintegration of ^{16}O and ^{63}Cu by Photons of Variable Energy. By J. G. Campbell	449
Photoprotons from Fluorine. By W. B. Lasich, E. G. Muirhead, and G. G. Shute	456
The Diffusion of Slow Electrons in Deuterium. By Barbara I. H. Hall ..	468
The Distribution of Radio Brightness over the Solar Disk at a Wavelength of 21 Centimetres. III. The Quiet Sun—Two-dimensional Observa- tions. By W. N. Christiansen and J. A. Warburton	474
Solar Brightness Distribution at a Wavelength of 60 Centimetres. I. The Quiet Sun. By G. Swarup and R. Parthasarathy	487
Apparent Intensity Variations of the Radio Source Hydra-A. By O. B. Slee	498
Rotation and Other Motions of the Magellanic Clouds from Radio Observa- tions. By F. J. Kerr and G. de Vaucouleurs	508
Movement of Sporadic E Ionization. By J. A. Harvey	523
Some Observations of Wind Velocity Autocorrelations in the Lowest Layers of the Atmosphere. By R. J. Taylor	535
Dispersal of Dust Particles from Elevated Sources. By G. T. Csanady ..	545

Short Communications

The Diffusion of Electrons in a Gas at Low Temperatures. By Barbara I. H. Hall	551
Lunar Tide in Sporadic E at Brisbane. By J. A. Thomas and A. C. Svenson	554
Electron Background in Imported G5 Emulsions. By V. D. Hopper and Jean E. Laby	557
An Unusual Electron Shower. By Y. K. Lim	562
The Angular Size of the Variable Radio Source Hydra-A. By A. W. L. Carter	564
Information on Energy Levels in ^8Be from Neutron Energy Spectra. By J. R. Bird and R. H. Spear	567
Corrigendum	573
Index to Volume 8	575

THE DISTRIBUTION OF DISLOCATIONS IN LINEAR ARRAYS

By A. K. HEAD* and N. LOUAT*

[Manuscript received October 8, 1954]

Summary

An approximate method is given for finding the equilibrium distribution of arrays of dislocations. The analysis is based on the assumption that an array of discrete dislocations may be replaced by a continuous distribution of smeared dislocation. The solutions of a number of problems of physical interest are investigated, including some in which dislocations of opposite sign are involved.

I. INTRODUCTION

Eshelby, Frank, and Nabarro (1951) have considered the problem of the position taken up by a set of identical straight dislocations, which are constrained to lie in some part of the same slip plane, under the combined action of their repulsions and the force exerted on them by a given applied shear stress.

Their method of solution, although exact, is subject to a number of disabilities in practice. In any particular problem it is apparently necessary to guess a function $q(n, x)$ which will give a polynomial solution to the differential equation representing the problem. If an appropriate $q(n, x)$ can be chosen then the dislocations will lie at the zeros of the polynomial solution. However, this polynomial may be one whose zeros are not tabulated and, if the number of dislocations, n , is large, much computation is then necessary. Furthermore, one does not usually obtain a general picture of dislocation distribution from such a computation.

We shall consider here an approximate method for the solution of the problem of Eshelby, Frank, and Nabarro. The approximation made is to replace the discrete distribution of finite dislocations by a continuous distribution of infinitesimal dislocations with the same total Burgers vector. The problem is then to find the density of dislocations at any point such that the distribution is in equilibrium under its mutual forces and that applied externally. This leads to the problem of inverting a singular integral equation, which is a routine procedure.

This approximation can be expected to involve little error when the distance between dislocations is of the same order as their width. Since these conditions are likely to occur near the head of any array of practical significance, it is also likely that the dislocation stresses near this region can be evaluated with reasonable accuracy if the equilibrium distribution of smeared-out dislocation can first be determined. One advantage of this method is that it will deal as easily with

* Aeronautical Research Laboratories, Department of Supply, Melbourne.

distributions of dislocation of opposite sign as with those of the same sign. The method of Eshelby, Frank, and Nabarro will not treat the case of the interaction of dislocations of opposite sign.

II. FORMULATION OF PROBLEM

Suppose that the plane $y=0$ is the slip plane, that the dislocation lines are parallel to the z -axis, and that they can move along the x -axis. Let $f(x)$ be the dislocation density at any point on the x -axis, with the convention that $f(x)$ be positive in regions of positive dislocations and vice versa. Let $P(x)$ be the appropriate component of the applied stress tending to move the dislocations along the x -axis, and take $P(x)$ to be positive if it tends to move a positive dislocation in the positive direction of the x -axis.

A small element of dislocation of strength ε/λ at x produces a stress at x_0 given by

$$\sigma = \frac{A\varepsilon}{x_0 - x}, \quad \dots\dots\dots (1)$$

where $A = \mu\lambda/2\pi$ for screw dislocations and $A = \mu\lambda/2\pi(1-\nu)$ for edge dislocations, where μ is the shear modulus of the material assumed isotropic, ν is Poisson's ratio, and λ is the Burgers vector of a unit dislocation. Hence the stress at x_0 due to the applied stress and the distributed dislocation is given by

$$S(x_0) = P(x_0) + A \int_D \frac{f(x)}{x_0 - x} dx + T(x_0), \quad \dots\dots\dots (2)$$

where the integral is taken over all regions D of the x -axis where there is dislocation present. $T(x)$ represents very short range stresses, with say the form of a Dirac δ function, which may be necessary for equilibrium at the ends of an array.

Since a dislocation will move if there is any net stress at its centre (excluding that produced by itself), the dislocation distribution can only be in equilibrium if $S(x_0)=0$ in the regions D of dislocation. Therefore we must have

$$\int_D \frac{f(x)}{x - x_0} dx - \frac{T(x_0)}{A} = -\frac{P(x_0)}{A} \quad \dots\dots\dots (3)$$

for all points x_0 in D and in particular for those at which $T(x)=0$. Since we exclude the self stress of a dislocation from the condition of equilibrium, the Cauchy principal value of the singular integral in (3) is to be taken. The inversion of this singular integral equation to find $f(x)$ for a given $P(x)$, with $T(x)=0$, is given by the following theorem.

III. INVERSION THEOREM

Singular integral equations, such as (3) have been investigated by Muskhelishvili (1953*a*, 1953*b*) and the inversion theorem for (3) is the following (Muskhelishvili 1953*a*, p. 251).

Suppose $P(x)$ is a known and $f(x)$ an unknown function and that D consists of p finite segments of the x -axis (a_1, b_1) , (a_2, b_2) , . . . , (a_p, b_p) . Suppose that

at q of the end-points of the segments, denoted by c_1, c_2, \dots, c_q , $f(x)$ is to remain bounded, and that at the remaining $2p-q$ end-points, denoted by $c_{q+1}, c_{q+2}, \dots, c_{2p}$, $f(x)$ may be unbounded. Let

$$R_1(x) = \prod_{k=1}^q (x - c_k), \quad R_2(x) = \prod_{k=q+1}^{2p} (x - c_k).$$

Then, if $p-q \geq 0$, solutions of (3), bounded at c_1, \dots, c_q , always exist and are given by

$$f(x_0) = \frac{-1}{\pi^2 A} \sqrt{\left\{ \frac{R_1(x_0)}{R_2(x_0)} \right\}} \int_D \sqrt{\left\{ \frac{R_2(x)}{R_1(x)} \right\}} \frac{P(x) dx}{x - x_0} + \sqrt{\left\{ \frac{R_1(x_0)}{R_2(x_0)} \right\}} Q_{p-q-1}(x_0), \quad \dots (4)$$

where $Q_{p-q-1}(x_0)$ is an arbitrary polynomial of degree not greater than $p-q-1$ (it is identically zero for $p=q$).

If $p-q < 0$, a unique solution, bounded at c_1, \dots, c_q , exists if and only if $P(x)$ satisfies the conditions

$$\int_D \sqrt{\left\{ \frac{R_2(x)}{R_1(x)} \right\}} x^n P(x) dx = 0, \quad \text{for } n=0, 1, \dots, (q-p-1), \dots (5)$$

and if this is so the solution is given by (4) with $Q_{p-q-1}(x) = 0$. Moreover at a bounded end-point, $f(x)$ vanishes.

IV. END CONDITIONS

Most of the problems considered by Eshelby, Frank, and Nabarro are of the type where a group of dislocations would move off to infinity owing either to their mutual repulsion or to an applied stress. They are prevented from doing so by a barrier in the form of a dislocation which is locked in position by a localized stress.

When setting up the same problem in our approximation we can proceed in one of two ways. Either we can leave the locked dislocation intact and only smear the free dislocations, or we can smear the locked dislocation too, in which case the barrier to the dislocation distribution becomes the localized stress field which locked the dislocation. This latter type of barrier, which we shall term a block, we take as a repulsive stress field which rises suddenly from zero to infinity. The effect of such a barrier has been taken into account in (2) and (3) by the term $T(x)$. For a locked dislocation it is found that the appropriate boundary condition on $f(x)$ is that it becomes zero at a small distance from the locked dislocation, this distance being a function of the stress forcing the distribution against the barrier.

V. EXAMPLES

The general method will be illustrated by some problems of physical interest. Some of these have already been considered by Eshelby, Frank, and Nabarro, and a comparison of the results obtained by the two methods indicates that the approximate method is little in error.

(i) *n positive dislocations in the potential trough given by $P(x) = -Cx$.* $f(x)$ will be symmetrical about $x=0$, and we assume it becomes zero at $x = \pm a$, where a will depend on n .

With the notation of Section III we have

$$\begin{aligned} p &= 1, & q &= 2, \\ R_1(x) &= (x^2 - a^2), & R_2(x) &= 1, \end{aligned}$$

and since $p - q < 0$ a solution will exist only if $P(x)$ satisfies (5). As this is so it follows that $f(x)$ is given by (4) to be

$$f(x) = \frac{C}{A\pi} \sqrt{(a^2 - x^2)}. \quad \dots\dots\dots (6)$$

The constant a is determined by

$$\int_{-a}^a f(x) dx = n,$$

which gives

$$a = \sqrt{(2nA/C)}. \quad \dots\dots\dots (7)$$

It is not possible to compare (6) with the corresponding result of Eshelby, Frank, and Nabarro, but they find that all dislocations lie in a region

$$|x| < \sqrt{\{(2n+1)A/C\}},$$

which is nearly equal to the value given by (7).

(ii) n positive dislocations between blocks at $x = \pm a$ with no applied shear stress. Since $P(x) = 0$ we get a non-zero solution for $f(x)$ only if we allow it to be unbounded at $x = \pm a$.

Then

$$\begin{aligned} p &= 1, \quad q = 0, \quad p - q > 0, \\ R_1(x) &= 1, \quad R_2(x) = x^2 - a^2, \\ f(x) &= Q_0 / \sqrt{(a^2 - x^2)}, \end{aligned}$$

where Q_0 is the arbitrary constant which the arbitrary polynomial Q_{p-q-1} in (4) becomes. Q_0 is determined by

$$n = \int_{-a}^a f(x) dx,$$

which gives

$$Q_0 = n/\pi.$$

(iii) n positive dislocations between unit positive dislocations locked at $x = \pm a$. Let $f(x)$ become zero at $x = \pm b$ ($b < a$). Then

$$\begin{aligned} p &= 1, \quad q = 2, \quad p - q < 0, \\ R_1(x) &= x^2 - b^2, \quad R_2(x) = 1, \end{aligned}$$

and a solution will exist if

$$P(x) = A \left\{ \frac{1}{x+a} + \frac{1}{x-a} \right\}$$

satisfies (5). This it does and (4) gives

$$f(x) = \frac{2a}{\pi \sqrt{(a^2 - b^2)}} \frac{\sqrt{(b^2 - x^2)}}{a^2 - x^2},$$

and b is determined by

$$n = \int_{-b}^b f(x) dx$$

to be

$$b = a \sqrt{\{1 - (\frac{1}{2}n + 1)^{-2}\}}.$$

(iv) n positive dislocations on the positive half of the x -axis, forced against a block at $x=0$ by a uniform stress $P(x) = -\sigma$. Let $f(x)$ be bounded at $x=a$ and unbounded at $x=0$. Then

$$p=1, \quad q=1, \quad p-q=0, \\ R_1(x) = x-a, \quad R_2(x) = x,$$

and (4) gives

$$f(x) = \frac{\sigma}{\pi A} \sqrt{\left(\frac{a-x}{x}\right)}. \quad \dots\dots\dots (8)$$

a is determined by

$$n = \int_0^a f(x) dx$$

as

$$a = 2nA/\sigma. \quad \dots\dots\dots (9)$$

Eshelby, Frank, and Nabarro have considered this case and (8) and (9) are in agreement with the expressions they have found by the exact treatment.

The total stress at any point on the x -axis due to the dislocation distribution and the applied stress is given by (2) as

$$S(x) = -\sigma \sqrt{\left(\frac{x-a}{x}\right)}, \quad x > a, \quad x < 0, \\ = 0, \quad a > x > 0.$$

(v) Blocks at $x = \pm a$ and a dislocation source at $x=0$. A uniform stress $P(x) = \sigma$ causes the source to generate equal numbers of positive and negative dislocations which move off in opposite directions until held up by the blocks. The source continues to generate dislocations until the net stress at the source is reduced to zero. $f(x)$ will be unbounded at $x = \pm a$, so

$$p=1, \quad q=0, \quad p-q > 0, \\ R_1(x) = 1, \quad R_2(x) = x^2 - a^2, \\ f(x) = \frac{\sigma}{\pi A} \frac{x}{\sqrt{(a^2 - x^2)}} + \frac{Q_0}{\sqrt{(a^2 - x^2)}},$$

where the arbitrary constant Q_0 is determined by the position of the source. Since $f(x)$ changes sign on either side of the source it will be zero at the position of the source. Hence Q_0 is zero for the simple symmetrical case.

The number of positive dislocations generated is given by

$$n = \int_0^a f(x) dx = \sigma a / \pi A, \quad \dots\dots\dots (10)$$

and is equal to the number of negative dislocations.

The stress on the x -axis beyond the blocks is given by (2) as

$$S(x) = \sigma \mid x \mid / \sqrt{(x^2 - a^2)}.$$

This is identical with the expression given by Starr (1928) for the stress in front of an infinitely narrow two dimensional crack extending from $x=-a$ to $x=a$ under a uniform applied stress. This is not surprising since it may be seen that the two problems are physically the same.

(vi) *Blocks at $x=\pm a$ with n positive dislocations between $x=b$ and $x=a$, and n negative dislocations, between $x=-b$ and $x=-a$, held apart by applied stress $P(x)=\sigma$. Let $f(x)$ be unbounded at $x=\pm a$ and bounded at $x=\pm b$. Then*

$$\begin{aligned} p=2, \quad q=2, \quad p-q=0, \\ R_1(x)=(x^2-b^2), \quad R_2(x)=(x^2-a^2), \\ f(x)=\pm \frac{\sigma}{\pi A} \sqrt{\left(\frac{x^2-b^2}{a^2-x^2}\right)}, \end{aligned}$$

where the sign is to be taken as positive for $b < x < a$ and as negative for $-a < x < -b$. b is related to the number of dislocation pairs by the relation

$$\begin{aligned} n &= \frac{\sigma}{\pi A} \int_b^a \sqrt{\left(\frac{x^2-b^2}{a^2-x^2}\right)} dx \\ &= \frac{\sigma}{\pi A} \left\{ aE[\sqrt{(1-b^2/a^2)}] - \frac{b^2}{a} K[\sqrt{(1-b^2/a^2)}] \right\}, \end{aligned}$$

where K and E are complete elliptic integrals of the first and second kinds respectively. If n is small, or to be more exact if b is nearly equal to a , this becomes

$$n \simeq \frac{\sigma}{2A} (a-b).$$

For b equal to zero this gives

$$n \simeq \sigma a / 2A,$$

which is approximately the value given by (10).

(vii) *n positive dislocations beyond $x=a$ and n negative dislocations beyond $x=-a$, driven together by a uniform stress $P(x)=-\sigma$, but prevented from coalescing by blocks at $x=\pm a$. Let the distribution be bounded at $x=\pm b$ and unbounded at $x=\pm a$. Then*

$$\begin{aligned} p=2, \quad q=2, \quad p-q=0, \\ R_1(x)=(x^2-b^2), \quad R_2(x)=(x^2-a^2), \\ f(x)=\pm \frac{\sigma}{\pi A} \sqrt{\left(\frac{b^2-x^2}{x^2-a^2}\right)}, \end{aligned}$$

and b is given by

$$\begin{aligned} n &= \frac{\sigma}{\pi A} \int_a^b \sqrt{\left(\frac{b^2-x^2}{x^2-a^2}\right)} dx \\ &= \frac{\sigma b}{\pi A} \{ K[\sqrt{(1-a^2/b^2)}] - E[\sqrt{(1-a^2/b^2)}] \} \\ &\simeq \frac{\sigma b}{\pi A} \left\{ \ln\left(\frac{4b}{a}\right) - 1 \right\}, \quad \text{for } b \gg a. \end{aligned}$$

VI. ACKNOWLEDGMENTS

The authors wish to thank the Chief Scientist, Department of Supply, for permission to publish this paper and Dr. L. H. Mitchell and Mr. J. R. M. Radok (the translator of the books of Muskhelishvili), of the Aeronautical Research Laboratories, for helpful discussions.

VII. REFERENCES

- ESHELBY, J. D., FRANK, F. C., and NABARRO, F. R. N. (1951).—*Phil. Mag.* **42**: 351.
MUSKHELISHVILI, N. I. (1953*a*).—"Singular Integral Equations." (P. Noordhoff, N.V.: Groningen.)
MUSKHELISHVILI, N. I. (1953*b*).—"Some Basic Problems of the Mathematical Theory of Elasticity." (P. Noordhoff, N.V.: Groningen.)
STARR, A. T. (1928).—*Proc. Camb. Phil. Soc.* **24**: 489.

NETS COMPOSED OF PARTS OF CIRCLES FOR THE APPROXIMATE SOLUTION OF FIELD PROBLEMS

By L. TASNY-TSCHIASSNY*

[Manuscript received July 15, 1954]

Summary

The two-dimensional differential equation

$$\frac{\partial}{\partial x} \left(\sigma \frac{\partial \varphi}{\partial x} \right) + \frac{\partial}{\partial y} \left(\sigma \frac{\partial \varphi}{\partial y} \right) + \tau = 0 \quad \dots \quad (1)$$

describes the current flow in a sheet of conductivity σ loaded by a transverse current density ($-\tau$), φ being the electrical potential. It is known that equation (1) can be solved approximately by a procedure in which the two-dimensional continuum is replaced by a net of straight-line bounded meshes, leading to an electrical network of conductances. The author shows that meshes bounded by "curvilinear rectangles" can be equally well dealt with and, on the basis of different conformal transformation functions for the individual meshes, derives the formulae required for a solution, if the mesh boundaries are circle arcs or circle arcs and straight lines. A good fit of the contours of the boundaries and equipotentials and their orthogonal trajectories can be obtained. This reduces the number of meshes without impairing the accuracy. Sharp corners at boundaries can be dealt with in a similar way. Formulae for a good accuracy computation of potential gradients and a method for changing the mesh size abruptly are given. Two examples using nets of only four meshes demonstrate the power of the method, the maximum errors being of the order of a few per cent.

I. INTRODUCTION

The problems dealt with in this paper are those governed by the two-dimensional differential equation

$$\frac{\partial}{\partial x} \left(\sigma \frac{\partial \varphi}{\partial x} \right) + \frac{\partial}{\partial y} \left(\sigma \frac{\partial \varphi}{\partial y} \right) + \tau = 0, \quad \dots \quad (1)$$

in which φ is an unknown function of position and σ and τ are known functions of positions or functions of φ and its derivatives, or both. An important problem of this type is the electric conduction in a plane sheet. It will be used for all explanations in this paper. Then φ is the electric potential, σ the electric conductivity, and τ the current density of external currents entering the sheet. In electrostatic field problems φ is the electric potential, σ is given by $1/(4\pi)$ -times the dielectric constant, and τ is the density of the space charge. Equation (1) covers also three-dimensional axially symmetrical arrangements if the distance from the axis (the radius) and the distance in the direction of the axis (the height) are dealt with as if they were Cartesian coordinates and the quantities substituted in equation (1) for σ and τ are the products of the radius and the actual values of σ and τ .

* Electrical Engineering Department, University of Sydney.

The approximate numerical solution of (1) by a system of simultaneous linear equations that can be considered the network equations of a system of conductances (and are amenable to relaxation methods) can be described in the following way :

Step 1.—Select a sufficient number of points within and on the boundaries for which the values of φ are to be found. Let these points be called nodes.

Step 2.—Find a linear relation

$$F(\varphi_N, \varphi_R, \varphi_S, \varphi_T, \dots) = \mathbf{R}_N \quad \dots\dots\dots (2)$$

between the value φ_N of φ at a node N and the values $\varphi_R, \varphi_S, \varphi_T, \dots$ at neighbouring nodes R, S, T, \dots which when complied with for $\mathbf{R}_N = 0$ secures that φ is an approximate solution of equation (1). Equation (2) takes the form of Kirchhoff's first rule with φ_N, φ_R , etc. denoting potentials. For an arrangement of the nodes as the corners of regular triangles, squares, and hexagons all conductances involved are equal, as Southwell (1946) has shown. The values of the conductances for an arrangement in which the nodes are the corners of irregular triangles can be found by formulae derived by the author (1949) and MacNeal (1953).

Step 3.—Solve the system of linear equations resulting from applying equation (2) to all nodes. A very convenient way of finding an approximate solution is Southwell's relaxation method (Motz and Worthy 1945 ; Southwell 1946 ; Tasny-Tschiasny 1949). The so-called residuals \mathbf{R}_N are computed for an arbitrarily selected set of values φ . A significant residual \mathbf{R}_N , usually the largest residual, is either liquidated or adjusted to a suitable value by changing the value of φ_N by a certain amount. By this the residuals $\mathbf{R}_R, \mathbf{R}_S, \mathbf{R}_T, \dots$ at the neighbouring nodes are altered too, but, in general, the changes are smaller than the change of the residual \mathbf{R}_N . Then another important residual is dealt with in the same way. The procedures converge fairly quickly and are continued until negligible values of all residuals are obtained. Instead of solving the system of linear equations numerically, analogues representing actual networks of conductances can be employed.

The boundaries require special artifices in the case of regular nets, because nodes need not necessarily lie on boundaries everywhere. If irregular nets are used, nodes can always be placed on the boundaries and no special problems arise. The errors in the values of φ , i.e. the differences between the values of φ , that comply with the system of linear equations (2) for $\mathbf{R}_N = 0$ and the values of φ that comply with the differential equation (1), are greater for irregular than for regular nets. For this reason and because the number of straight lines simulating a sharply curved part of the boundary must be large, the number of nodes must also be large. This increases the labour in solving the simultaneous equations.

In the present paper we introduce the use of those curvilinear nets in which the mesh contours are parts of circles or parts of circles and straight lines. Basing our derivations on the conformal transformation of a curvilinear into a rectilinear mesh, in Section II methods are developed by which the interior of a "curvilinear rectangle" can be approximated by lumped conductances connected between its corners. This approximation permits the use of different

transformation functions for the different meshes of a curvilinear net as long as these functions supply the same curve for the common boundary of adjacent meshes. If a suitable net is laid out and the interior of all meshes replaced by the conductances mentioned, an electrical network results in which the statement of Kirchhoff's first rule supplies the required equation (2).

The error occurring when using the described nets appropriately is much smaller than the error involved in nets with straight contours. The additional labour spent in laying out a curvilinear net may be often compensated for by the smaller number of nodes required for the same accuracy. Since it can always be arranged that nodes are on the boundaries, as in a net formed of irregular triangles, no special artifices are required for the boundaries. Contours used in engineering are often composed of parts of circles and straight lines; hence the shape of the boundary can generally be exactly adhered to. In certain types of problems, for instance, the problem of finding the maximum value of the voltage gradient occurring in a material, this is an advantage, because the maximum voltage gradient occurs usually at the boundaries.

II. THE REPLACEMENT OF THE INTERIOR OF A "CURVILINEAR RECTANGLE" BY LUMPED CONDUCTANCES

Let

$$w = u + jv = w(z) = w(x + jy) \quad \dots \dots \dots (3)$$

be an analytical function. Then u and v comply with the Cauchy-Riemann differential equations

$$\left. \begin{aligned} \frac{\partial u}{\partial x} &= \frac{\partial v}{\partial y}, \\ \frac{\partial u}{\partial y} &= -\frac{\partial v}{\partial x}. \end{aligned} \right\} \dots \dots \dots (4)$$

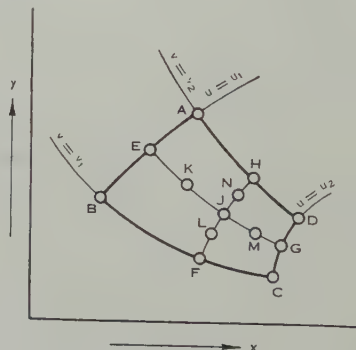
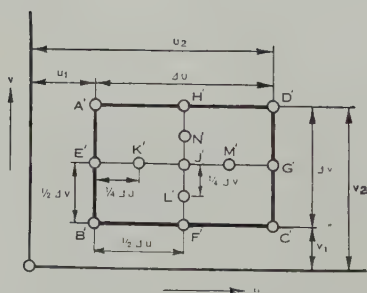


Fig. 1.—Rectilinear rectangle in the w -plane.

Fig. 2.—Curvilinear rectangle in the z -plane.

A rectangle $A'B'C'D'$ in the w -plane (Fig. 1) with the mid points E' , F' , G' , H' of its sides, its centre J' , and the mid points K' , L' , M' , and N' between J' and E' , F' , G' , and H' respectively is the result of a transformation by $w(z)$ of a "curvilinear rectangle" $ABCD$ in the z -plane (Fig. 2) marked correspondingly without primes. Let x and y be the variables appearing in equation (1) and let the curvilinear rectangle $ABCD$ (Fig. 2) be a mesh of a curvilinear net with the

nodes A , B , C , and D . To replace the interior of this mesh by a network of conductances we proceed in the following way. First we express the difference in potential $(\varphi_A - \varphi_B)$ between the points A and B as an integral taken along the contour BEA , i.e. along the contour defined by $u = u_1$.

$$\varphi_A - \varphi_B = \int_B^A \frac{\partial \varphi}{\partial v} dv. \quad \dots\dots\dots (5)$$

If $\partial \varphi / \partial v$ is expressed by Taylor's expansion about the point E and the integration carried out we obtain

$$\varphi_A - \varphi_B = \left(\frac{\partial \varphi}{\partial v} \right)_E \Delta v + O(\Delta v^3), \quad \dots\dots\dots (6)$$

where the subscript E denotes the value at the point E and Δv is given by

$$\Delta v = v_2 - v_1. \quad \dots\dots\dots (7)$$

The term $O(\Delta v^3)$ contains $(\partial^3 \varphi / \partial v^3)_E$ and higher derivatives of φ .

In the approximation by lumped conductances the current passing within the conducting sheet through the line $EKJLF$ is to be made equal to the current collected at the point B . The current I_{EJ} through the line EKJ is given by

$$I_{EJ} = \int_E^J \sigma \left(\frac{\partial \varphi}{\partial y} dx - \frac{\partial \varphi}{\partial x} dy \right). \quad \dots\dots\dots (8)$$

If $\partial \varphi / \partial x$ and $\partial \varphi / \partial y$ are expressed in terms of $\partial \varphi / \partial u$, $\partial \varphi / \partial v$, $\partial u / \partial x$, $\partial v / \partial x$, $\partial u / \partial y$, and $\partial v / \partial y$, and equations (4) are used, expressions for the total differentials du and dv result. The formula

$$I_{EJ} = \int_E^J \sigma \left(\frac{\partial \varphi}{\partial v} du - \frac{\partial \varphi}{\partial u} dv \right) \quad \dots\dots\dots (9)$$

is obtained. For the contour EKJ dv is zero and the second term in the bracket vanishes. With the aid of various expansions according to Taylor's theorem this integral can be approximated on the basis of σ_K , $(\partial \varphi / \partial v)_E$, and $(\partial^2 \varphi / \partial u \partial v)_J$, where the subscripts denote the values at the appropriate points. If, further, $(\partial^2 \varphi / \partial u \partial v)_J$ is approximated by

$$\left(\frac{\partial^2 \varphi}{\partial u \partial v} \right)_J = \frac{\varphi_B + \varphi_D - \varphi_A - \varphi_C}{\Delta u \Delta v} + O(\Delta v^2) + O(\Delta u^2), \quad \dots\dots (10)$$

and the resulting expression for I_{EJ} divided by equation (6) we obtain

$$\frac{I_{EJ}}{\varphi_A - \varphi_B} = \sigma_K \frac{\Delta u}{2 \Delta v} \left[1 + \frac{1}{4} \frac{\varphi_B + \varphi_D - \varphi_A - \varphi_C}{\varphi_A - \varphi_B} + O(\Delta u^2) + O(\Delta v^2) \right]. \quad \dots (11)$$

The terms $O(\Delta u^2)$ and $O(\Delta v^2)$ contain expressions in σ and φ obtained by at least three differentiations, with respect to u or v . The term $\frac{1}{4}(\varphi_B + \varphi_D - \varphi_A - \varphi_C) / (\varphi_A - \varphi_B)$ is $O(\Delta u)$, as can be seen from equations (6) and (10).

Expressions similar to equation (11) can be obtained for $I_{FJ}/(\varphi_C - \varphi_B)$, $I_{GJ}/(\varphi_D - \varphi_C)$, and $I_{HJ}/(\varphi_D - \varphi_A)$. Closer scrutinizing, i.e. computing $(I_{EJ} + I_{FJ})$, $(I_{FJ} + I_{GJ})$, $(I_{GJ} + I_{HJ})$, and $(I_{HJ} + I_{EJ})$, results in two alternative networks of conductances replacing the interior of the curvilinear rectangle $ABCD$. The first alternative (see Fig. 3) neglects errors $O(\Delta u)$ and $O(\Delta v)$. This means that

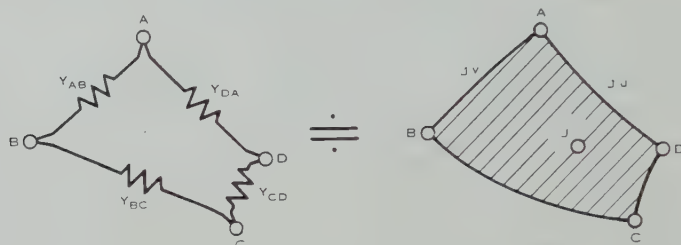


Fig. 3.—Approximate representation of a mesh with a relative error of the order of Δu and Δv .

$$Y_{AB} = Y_{CD} = \frac{1}{2} \sigma_J \left| \frac{\Delta u}{\Delta v} \right|,$$

$$Y_{DA} = Y_{BC} = \frac{1}{2} \sigma_J \left| \frac{\Delta v}{\Delta u} \right|.$$

terms like $\frac{1}{4}(\varphi_B + \varphi_D - \varphi_A - \varphi_C)/(\varphi_A - \varphi_B)$ in equation (11) are neglected and that σ_K , σ_L , σ_M , and σ_N can be replaced by σ_J . Figure 3 gives the details of the network. In the second alternative (Fig. 4) the error is $O(\Delta u^2)$ plus $O(\Delta v^2)$. Terms like $\frac{1}{4}(\varphi_B + \varphi_D - \varphi_A - \varphi_C)/(\varphi_A - \varphi_B)$ in equation (11) are retained, but it is

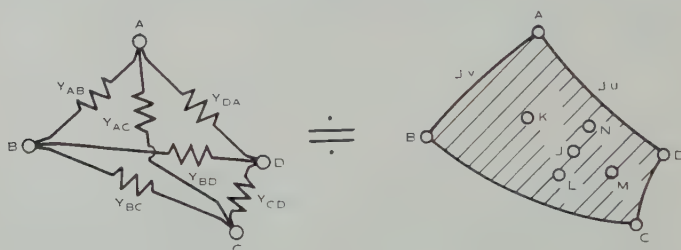


Fig. 4.—Approximate representation of a mesh with a relative error of the order of Δu^2 and Δv^2 .

$$Y_{AB} = \frac{1}{2} \sigma_K \left| \frac{\Delta u}{\Delta v} \right| - Y_{AC}; \quad Y_{CD} = \frac{1}{2} \sigma_M \left| \frac{\Delta u}{\Delta v} \right| - Y_{AC};$$

$$Y_{BC} = \frac{1}{2} \sigma_L \left| \frac{\Delta v}{\Delta u} \right| - Y_{AC}; \quad Y_{DA} = \frac{1}{2} \sigma_N \left| \frac{\Delta v}{\Delta u} \right| - Y_{AC};$$

$$Y_{AC} = Y_{BD} = \frac{1}{2} \sigma_J \left(\left| \frac{\Delta u}{\Delta v} \right| + \left| \frac{\Delta v}{\Delta u} \right| \right).$$

admissible to replace the multiplying factors σ_K , σ_L , σ_M , and σ_N by σ_J as far as these terms are concerned. In practice one will usually replace σ_K , σ_L , σ_M , and σ_N by σ_J throughout or by values pertaining to one of the corners of the curvilinear square $ABCD$, because the variation of σ with position will not be rapid.

For $\sigma = \text{constant}$ the error vanishes if the lines of constant φ in the w -plane are straight, because then all derivatives of φ with respect to u and v higher than the first vanish and these higher derivatives are multiplying factors in the terms $O(\Delta u^2)$ and $O(\Delta v^2)$ of equations (10) and (11). In particular this is the case if two opposite sides of the curvilinear rectangle $ABCD$ coincide with lines of constant potential.

The method described requires that a loading of the curvilinear rectangle $BFJE$ by external currents is lumped as an external current applied at the node B . Similarly the loadings of the rectangles $FCGJ$, $JGDH$, and $EJHA$ are concentrated at the nodes C , D , and A respectively. The magnitudes of the concentrated currents can be computed approximately as the products of the areas of the rectangles concerned and mean values of the specific loading ($-\tau$).

III. NETS IN WHICH THE MESH CONTOURS ARE PARTS OF CIRCLES AND STRAIGHT LINES

The results of Section II show that there is no objection to employing different analytical functions for different meshes of the net as long as adjacent contours coincide. In this section it will be shown how curvilinear rectangles bounded by parts of circles or straight lines can be conveniently dealt with.

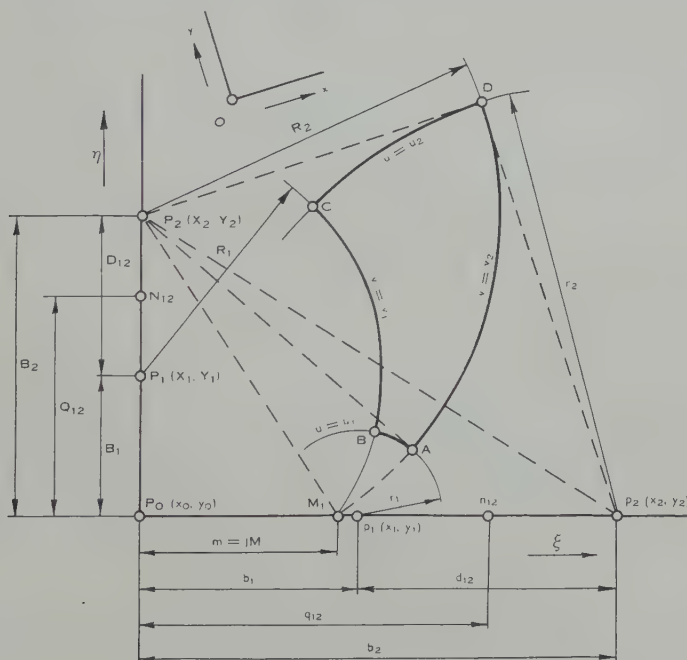


Fig. 5.—Curvilinear rectangle composed of arcs of circles.

(a) The Field Produced by One Source and One Sink

In a system of Cartesian coordinates ξ, η (Fig. 5) that does not usually coincide with the system of coordinates x, y used in equation (1), let the point $M_1(m, 0)$ be a sink and the point $M_2(-m, 0)$, not shown in the diagram, be a

source of current, both of the same intensity. The ξ -axis will be called the source axis, the η -axis the sourceless axis, and the origin P_0 the geometric centre. If the intensity of source and sink is appropriately chosen, the analytical function

$$w = \ln \frac{\xi - m}{\xi + m} \quad \dots\dots\dots (12)$$

of $\zeta = \xi + j\eta$ supplies in its real part

$$u = \frac{1}{2} \ln \frac{(\xi - m)^2 + \eta^2}{(\xi + m)^2 + \eta^2} \quad \dots\dots\dots (13)$$

the family of equipotentials (u is the parameter of the family), and in its imaginary part

$$v = \tan^{-1} \frac{\eta}{\xi - m} - \tan^{-1} \frac{\eta}{\xi + m} \quad \dots\dots\dots (14)$$

the family of flow functions (v is the parameter of the family) peculiar to this arrangement. By appropriate manipulations on equations (13) and (14) or by straight-out verification it can be shown that a circle with centre $(b_1, 0)$ and radius r_1 where

$$r_1^2 = b_1^2 - m^2 \quad \dots\dots\dots (15)$$

is an equipotential for a value of the potential

$$u_1 = \sinh^{-1} \left(\frac{m}{r_1} \right) \quad \dots\dots\dots (16)$$

(and similarly for other subscripts) and that a circle with centre (O, B_1) and radius R_1 where

$$R_1^2 = B_1^2 + m^2 \quad \dots\dots\dots (17)$$

is a flow line for the value of the flow function

$$v_1 = \sin^{-1} \left(\frac{m}{R_1} \right) \quad \dots\dots\dots (18)$$

(and similarly for other subscripts). If the quantity m , henceforth called the parameter, is given and two pairs of values u_1, u_2 and v_1, v_2 are selected, four circles result from equations (15)-(18) and determine a curvilinear rectangle $ABCD$ (Fig. 5). This rectangle can be replaced by a network of conductances (see Figs. 3 and 4). The quantities Δu and Δv are the absolute values of the differences $(u_1 - u_2)$ and $(v_1 - v_2)$ respectively.

Equations (15) and (17) on the one hand and equations (16) and (18) on the other become identical, if fictitious quantities

$$\left. \begin{aligned} U &= jv, \\ M &= jm \end{aligned} \right\} \quad \dots\dots\dots (19)$$

are introduced and in writing down the equations *either* capital letters *or* small letters are used. This symmetry with small letters referring to the u -circles and capital letters referring to the v -circles is useful.

(b) *Basic Relations for a Curvilinear Rectangle*

In Figure 5 besides the axes of coordinates ξ and η with the origin (the geometric centre) P_0 the axes of coordinates x and y , referring to equation (1) are shown. The coordinates added in parentheses to the individual points refer to the frame (x, y) . In the curvilinear rectangle $ABCD$ p_1, p_2 are the centres of the u -circles $u=u_1, u=u_2$, and P_1, P_2 the centres of the v -circles $v=v_1, v=v_2$. The points n_{12} and N_{12} are the mid points between the respective circle centres. From Pythagoras's theorem applied to the triangles P_2p_2D and $P_2p_2P_0$ we obtain

$$\overline{P_2D^2} = \overline{P_2P_0^2} + \left(q_{12} + \frac{d_{12}}{2} \right)^2 - r_2^2, \quad \dots\dots\dots (20)$$

where d_{12} is the distance between the centres of the two u -circles and q_{12} the distance between n_{12} and P_0 . Similarly we obtain

$$\overline{P_2A^2} = \overline{P_2P_0^2} + \left(q_{12} - \frac{d_{12}}{2} \right)^2 - r_1^2. \quad \dots\dots\dots (21)$$

Since $\overline{P_2D} = \overline{P_2A}$ we obtain from equations (20) and (21)

$$q_{12} = \frac{r_2^2 - r_1^2}{2d_{12}}. \quad \dots\dots\dots (22)$$

It follows from the triangle $M_1P_0P_2$ that

$$\overline{P_2M_1^2} = \overline{P_2P_0^2} + m^2. \quad \dots\dots\dots (23)$$

Since $\overline{P_2D} = \overline{P_2A} = \overline{P_2M_1}$ and

$$\left. \begin{aligned} b_1 &= q_{12} - \frac{d_{12}}{2}, \\ b_2 &= q_{12} + \frac{d_{12}}{2}, \end{aligned} \right\} \quad \dots\dots\dots (24)$$

we obtain from equations (20), (21), and (23) the conditions for orthogonality (see equation (15))

$$\left. \begin{aligned} m^2 &= b_2^2 - r_2^2, \\ m^2 &= b_1^2 - r_1^2, \end{aligned} \right\} \quad \dots\dots\dots (25)$$

and

$$m^2 = q_{12}^2 + \left(\frac{d_{12}}{2} \right)^2 - \frac{r_2^2 + r_1^2}{2}. \quad \dots\dots\dots (26)$$

Equations (22), (24), (25), and (26) refer to the u -circles. It can be easily shown that these equations hold good for the v -circles, if the quantities $r_1, r_2, q_{12}, d_{12}, b_1, b_2$, and m are replaced by the corresponding capital letter quantities (see Fig. 5 and equation (19)).

In Section IV it will be discussed in detail how these relations can be utilized to solve the problems connected with the layout of a curvilinear net. At present it should be pointed out only that the frame (ξ, η) and the value of m are unequivocally determined if two u -circles or two v -circles are given.

(c) *General Points Regarding the Layout of a Net*

It is advisable to work with the same frame (x, y) for the whole net. Figure 6 shows a convenient way of recording in a single figure for the whole net: the centres (x_1, y_1) , (x_2, y_2) , (x_3, y_3) , (x_4, y_4) of four contour circles, their radii r_1, r_2, r_3, r_4 , the geometric centre (x_0, y_0) , and the parameter m of a mesh $ABCD$ and, with the aid of the arrowed lines starting at the value of m , which circles are the v -circles. If a circle degenerates into a straight line (e.g., the line AF , Fig. 6), the direction tangent a_6 of the line and a point (x_6, y_6) through which it passes are indicated, instead of the position of the centre and the length of the radius which are infinite.

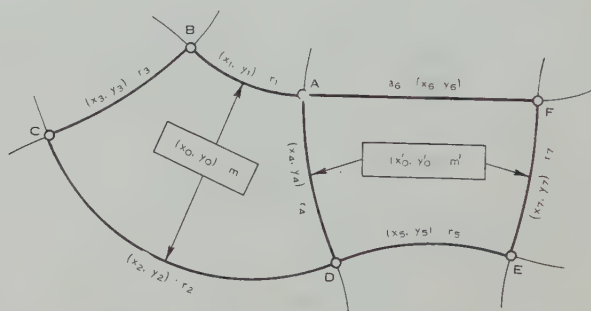


Fig. 6.—Method of recording the characteristic data in a diagram.

If a net consisting entirely of curvilinear rectangles is to be laid out, one tries to follow approximately the direction of the equipotentials and their orthogonal trajectories. If curvilinear meshes are to be employed near the boundaries only, one tries to avoid too abrupt changes of the angle between the directions of the equipotentials and the mesh boundaries. These procedures ensure that the lines of constant φ in the w -plane are only slightly curved (see end of Section II). After forming an idea of the mesh sizes in the individual parts of the field one starts at a boundary and proceeds from mesh to mesh. Thereby problems 1, 2, and 3, to be dealt with in Section IV (a), are to be solved in succession. Meshes in which simultaneously two v -circles degenerate into straight lines are dealt with in Section IV (b). If the mesh size is to be changed, the method described in Section VI is used. Sharp corners occurring at the boundaries or at the surfaces between dielectrics of different dielectric constants can be included by a procedure described in Section V.

IV. DETAILED PROCEDURES FOR THE LAYOUT OF A NET

If the accuracy requirements are not very stringent it will suffice for most of the meshes to rely on the drawing and to measure the required dimensions. For some meshes, or if greater accuracy is required, for a considerable number of

them computations must replace measurements. It is recommended that computations be carried out in Cartesian coordinates common to all meshes, as mentioned before. The procedures most suitable for use with Cartesian coordinates are described below.

(a) *At least One of the u -Circles and One of the v -Circles do not degenerate into a Straight Line*

Problem 1.—Two u -circles or two v -circles are given. It is not necessarily known whether they are u - or v -circles. Find the geometric centre and the value of the parameter.

Solution.—Refer to Figure 5. Since it is not known whether the circles are u - or v -circles, small letters will be used for the symbols, but the procedure is similar for capital letter symbols.

- (1) Find the mid point n_{12} between the two centres of the circles.
- (2) Using equation (22) find the length q_{12} .
- (3) On the line joining the centres of the two circles transfer the length q_{12} from n_{12} to that side on which the centre of the smaller circle lies. This determines the geometric centre P_0 .
- (4) Using equation (24) compute b_1 or b_2 and using equation (25) compute m . Alternatively, m can be found directly from equation (26). If m is real, the two given circles are u -circles, if m is imaginary, they are v -circles.

If one of the given circles degenerates into a straight line, the geometric centre is found as the intersection of this line with the perpendicular to it through the centre of the non-degenerate circle. The value of the parameter is given by one of the two equations (25).

Problem 2.—The geometric centre, the source and sourceless axes, the value of the parameter, and a point are given. Find the u -circle and the v -circle passing through the given point.

Solution.—Refer to Figure 5. Let D be the given point. If either p_2 , the centre of the u -circle through D , or P_2 , the centre of the v -circle through D are given, draw a perpendicular to the line p_2D (or P_2D) through D and intersect with the axis on which $p_2(P_2)$ does not lie. The point of intersection is the centre of the circle not given. If neither circle through D is given, find the centre of the v -circle through D as the intersection of the bisector of DM_1 (or DM_2) and the sourceless axis.

Problem 3.—One u - and one v -circle, the geometric centre, the two axes, and the value of the parameter are given. Find the two points of intersection, one of which will be used.

Solution.—Refer to Figure 5. Let the given circles be the u_2 - and v_2 -circle with the centres p_2 and P_2 respectively. The graphical solution is straightforward. Analytically the point D can, often with less labour, be found as the point of intersection of the straight lines p_2D and P_2D . If a is the direction tangent of the line p_2P_2 , the direction tangents of the lines p_2D and P_2D equal $\tan [\tan^{-1} a \pm \tan^{-1} (R_2/r_2)]$ and $\tan [\tan^{-1} a \mp \tan^{-1} (r_2/R_2)]$ respectively.

If one of the two circles degenerates into a straight line, let it be called the circle 1. Then the distance from the geometric centre P_0 of the points of intersection between the other circle 2 and this straight line is equal to $(b_2 \pm r_2)$ or $(B_2 \pm R_2)$, as the case may be, positive values being on the side of P_0 on which p_2 or P_2 lies.

b_2 or B_2 is to be computed from equation (25).

(b) *Both v -Circles Degenerate into Straight Lines*

In this case the bundle of v -circles degenerates into a pencil of straight lines and the u -circles are concentric circles with their centre in the point of intersection of the v -lines. If the direction tangents of two v -lines are a_1 and a_2 and the radii of two u -circles R_1 and R_2 , we obtain

$$\Delta u = u_2 - u_1 = \ln \frac{R_2}{R_1} = 2 \cdot 30259 \log_{10} \left(\frac{R_2}{R_1} \right), \quad \dots \dots \dots (27)$$

$$\Delta v = v_2 - v_1 = \tan^{-1} a_2 - \tan^{-1} a_1 = \tan^{-1} \frac{a_2 - a_1}{1 + a_1 a_2}. \quad \dots (28)$$

(c) *Notes Regarding the Computation of Δu and Δv*

When using equations (16) and (18) for the computation of u_1, u_2, v_1, v_2 care must be exercised—because both \sinh^{-1} and \sin^{-1} are multivalued functions. The following rules eliminate any possibility of an error in the computation of $|\Delta u / \Delta v|$ and $|\Delta v / \Delta u|$, that is, the quantities required for the computation of the conductances in Figures 3 and 4.

Rule for the Computation of Δu

To find $|\Delta u|$ take the difference of $|u_1|$ and $|u_2|$ if the sourceless line is outside the curvilinear square, and add $|u_1|$ and $|u_2|$ if it passes through it.

Rules for the Computation of Δv

- (1) *Definition of "small" and "great" arcs.*—Let that part of the v -circle that lies between the sources M_1 and M_2 and contains the arc considered be drawn (or thought to be drawn). If this part of the circle is greater than a half-circle, viz. if the centre of the circle is within the area defined by the part of the v -circle drawn and the straight line connecting the sources M_1 and M_2 , the arc shall be called a "great" arc. If this is not the case the arc shall be called a "small" arc.
- (2) If $\sin^{-1}(m/R_1)$ is the value of $\sin^{-1}(m/R)$ that is between 0 and $\frac{1}{2}\pi$, then

$$v_1 = \sin^{-1} \left(\frac{m}{R_1} \right) = \sin^{-1} \left(\frac{m}{R_1} \right), \quad \text{for "small" arcs,}$$

$$v_1 = \sin^{-1} \left(\frac{m}{R_1} \right) = \pi - \sin^{-1} \left(\frac{m}{R_1} \right), \quad \text{for "great" arcs.}$$

- (3) To find $|\Delta v|$, take the difference of $|v_1|$ and $|v_2|$ if the source line is outside the curvilinear square, and add $|v_1|$ and $|v_2|$ if it passes through it.

V. SHARP CORNERS

Sharp corners may occur at the electrodes and at the border lines of different dielectrics. Usually a sharp corner is formed by two straight lines. If this is not the case it can for a certain distance be approximated by a corner of this type to make the following treatment possible.

In Figure 7 let aOc be the sharp corner of the aperture

$$\alpha = p\pi, \quad \dots \dots \dots (29)$$

and let Ob be the bisector of α . Let the distance $OA=OB=OC=t$ be conveniently chosen. Let three circles of equal radius R with their centres on the lines Oa , Ob , and Oc be drawn in such a way that they intersect at right angles

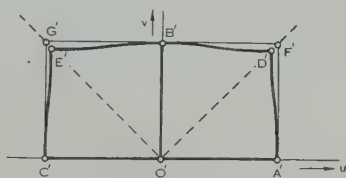
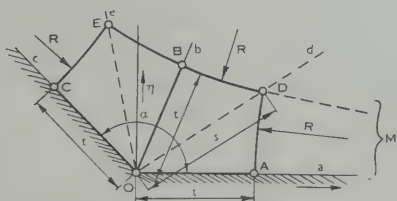


Fig. 7.—Sharp corner in the ζ -plane. Fig. 8.—Sharp corner in the w -plane.

at the points D and E that lie on the bisectors Od and Oe of the angles aOb and bOc respectively. The analysis of the triangle ODM shows that the radii R of these circles and the distances s of the points D and E from the corner O are given by

$$R = \frac{t \cdot \sin(p\pi/4)}{1/\sqrt{2} - \sin(p\pi/4)}, \quad \dots \dots \dots (30)$$

$$s = \frac{t \cdot \sin[(1-p)\pi/4]}{1/\sqrt{2} - \sin(p\pi/4)}, \quad \dots \dots \dots (31)$$

The ratios (R/t) and (s/t) for a few typical angles α are contained in Table 1. If a system of Cartesian coordinates (ξ, η) with the origin O and the direction of the ξ -axis coinciding with the direction Oa is chosen (Fig. 7), the configuration of Figure 7 can be conformally transformed into the configuration of Figure 8 (Schwarz-Christoffel transformation) and the transforming function is

$$\zeta = \xi + j\eta = t \left(\frac{w}{t} \right)^p = t \left(\frac{u + jv}{t} \right)^p. \quad \dots \dots \dots (32)$$

The curvilinear square $O'A'D'B'$ deviates only slightly from the rectilinear square $O'A'F'B'$. This is evident from Table 1, in which the ratios

$$\frac{O'D'}{O'F'} = \left(\frac{s}{t} \right)^{1/p} \cdot \frac{1}{\sqrt{2}} \quad \dots \dots \dots (33)$$

are tabulated for various angles α .

TABLE 1
RATIOS (R/t) , (s/t) , and $(O'D'/O'F')$ FOR TYPICAL ANGLES α

α	45°	90°	135°	180°	225°	270°	315°
p	0.250	0.500	0.750	1.000	1.250	1.500	1.750
R/t	0.381	1.180	3.667	∞	-6.684	-4.262	-3.597
s/t	1.085	1.178	1.288	1.414	1.570	1.762	2.035
$O'D'/O'F'$..	0.980	0.982	0.991	1.000	1.014	1.031	1.061

An argument based on the subdivision of the figures $OADB$ and $OBEC$ into (n^2-1) parts which are nearly curvilinear squares (n is an integer) and one figure which is geometrically similar to the original figure, shows that the figures $OADB$ and $OBEC$ can in very good approximation be replaced by the networks of conductances shown in Figures 3 and 4, as if they were curvilinear squares ($\Delta u = \Delta v = t$) with the corners at O, A, D, B and O, B, E, C respectively.

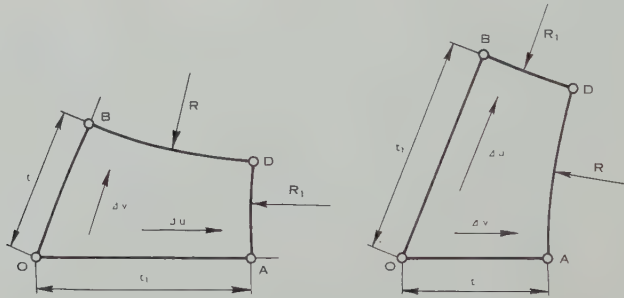


Fig. 9.—Curvilinear rectangle with a sharp corner, $t_1 > t$.

It is sometimes convenient to make the lengths OA and OB slightly different. In this case Figures 3 and 4 can still be used, if the following specifications referring to Figure 9 for $t_1 > t$ are adhered to, in which the directions of Δu and Δv are indicated.

R is given by equation (30).

$$R_1 = \frac{(t-t_1)[(t-t_1)/\sqrt{2} + (t+t_1) \sin(p\pi/4)] + 2\sqrt{2} t t_1 \sin^2(p\pi/4)}{2[1/\sqrt{2} - \sin(p\pi/4)] \cdot [t \cdot \sqrt{2} \cdot \sin(p\pi/4) + (t-t_1)]}, \quad (34)$$

$$\left. \begin{aligned} \frac{\Delta u}{\Delta v} &= \left(\frac{t_1}{t}\right)^{1/p}, \\ \frac{\Delta v}{\Delta u} &= \left(\frac{t}{t_1}\right)^{1/p}. \end{aligned} \right\} \quad (35)$$

Equation (34) is derived from the condition that for given values of t and t_1 and for R given by equation (30), the circles of radii R and R_1 intersect at right angles. Equations (35) are the consequence of the transformation equation (32).

VI. CHANGE OF THE MESH SIZE

In parts of the field in which the field gradient is smaller and does not change rapidly an increase of the mesh size reduces the labour considerably without affecting the degree of accuracy. If in Figure 10 the circle which passes through the points A , B , and C forms the boundary AC in the curvilinear rectangle $ACDE$ and the boundary CB in the curvilinear rectangle $CBFD$ —which can be arranged for in the layout of the net—the node C can be eliminated in the

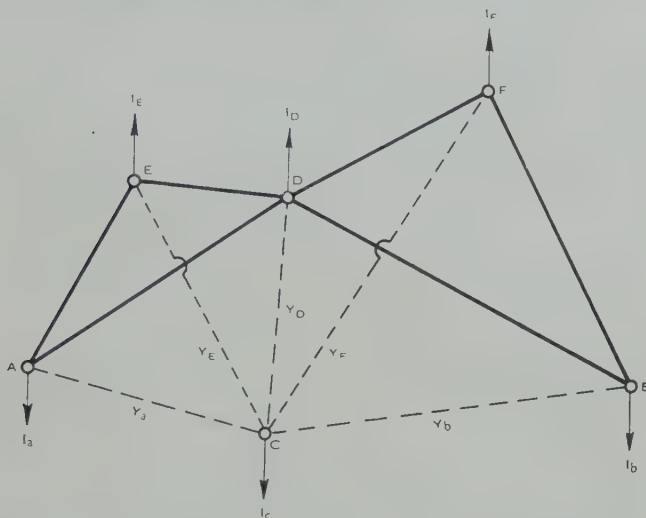


Fig. 10.—Part of a net with a node C not yet eliminated.

following way. Prerequisites are that, in the neighbourhood of C , σ does not vary very much and the equipotentials and field lines in the w -plane are nearly straight. Then we can assume that approximately

$$\varphi_c = k_a \varphi_a + k_b \varphi_b, \quad \dots \dots \dots (36)$$

where

$$\left. \begin{aligned} k_a &= \frac{Y_a}{Y_a + Y_b}, \\ k_b &= \frac{Y_b}{Y_a + Y_b}, \end{aligned} \right\} \dots \dots \dots (37)$$

and where Y_a and Y_b are the conductances connecting the nodes CA and CB respectively. Let us split each of the conductances Y_E , Y_D , and Y_F that connect the nodes E , D , and F respectively and C (Fig. 10) into two parallel conductances $(k_a Y_E)$, $(k_a Y_D)$, $(k_a Y_F)$ leading to node A , and $(k_b Y_E)$, $(k_b Y_D)$, $(k_b Y_F)$ leading to node B respectively (Fig. 11). Let us, further, connect the nodes A and B by a conductance equal to the series combination of Y_a and Y_b . The broken lines in Figures 10 and 11 are the conductances that are affected by this procedure and the full lines are those that are not. If the potentials φ_A , φ_B , φ_D , φ_E , φ_F of the nodes A , B , D , E , and F , and the external node currents I_A , I_B , I_D , I_E , and I_F at these nodes (Figs. 10 and 11) respectively are assumed to be equal in pairs,

the analyses of Figures 10 and 11 show that the current I_C appearing in Figure 10 is split into two parts ($k_a I_C$) and ($k_b I_C$) loading additionally the nodes A and B as shown in Figure 11. Since this way of accounting for the current I_C is reasonable, the given method for the elimination of the node C is sound.

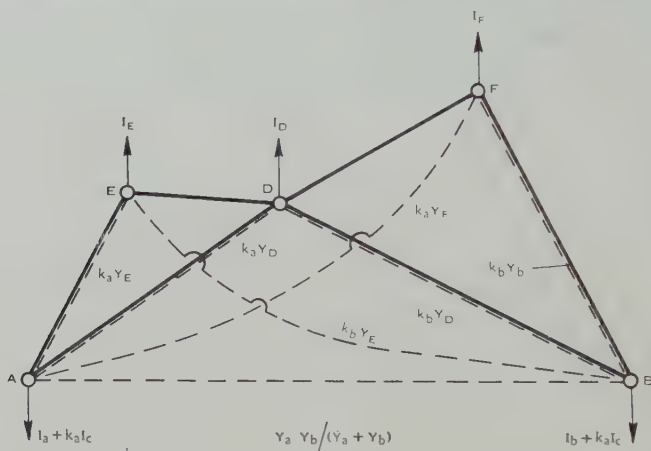


Fig. 11.—Part of the net of Figure 10 with the node C eliminated.

VII. THE COMPUTATION OF THE POTENTIAL GRADIENTS

(a) Curvilinear Squares

If the values φ_A , φ_B , φ_C , and φ_D of the potentials at the four corners of a curvilinear rectangle $ABCD$ (Fig. 2) are given, the potential gradient at any point that is not outside this rectangle can be computed with the aid of the formulae

$$\left. \begin{aligned} g_u &= \frac{\partial \varphi}{\partial u} \cdot \left| \frac{dw}{d\zeta} \right|, \\ g_v &= \frac{\partial \varphi}{\partial v} \cdot \left| \frac{dw}{d\zeta} \right|. \end{aligned} \right\} \dots \dots \dots (38)$$

In equations (38) g_u and g_v are the potential gradients in the directions of the orthogonal lines $v=\text{constant}$ and $u=\text{constant}$ respectively, so that the magnitude g of the gradient is given by

$$g^2 = g_u^2 + g_v^2. \dots \dots \dots (39)$$

If the field produced by φ in the w -plane is nearly uniform, the value $\partial \varphi / \partial u$ can be interpolated from the values $(\partial \varphi / \partial u)_{AD}$ and $(\partial \varphi / \partial u)_{BC}$ where $(\partial \varphi / \partial u)_{AD}$ can be approximated by

$$\left(\frac{\partial \varphi}{\partial u} \right)_{AD} = \frac{\varphi_D - \varphi_A}{\Delta u}, \dots \dots \dots (40)$$

and similarly for $(\partial \varphi / \partial u)_{BC}$ and $\partial \varphi / \partial v$. For $|dw/d\zeta|$ we find by differentiation of (12) after some manipulations

$$\left| \frac{dw}{d\zeta} \right| = \frac{2m}{\sqrt{\{(\rho^2 - m^2)^2 + 4m^2\eta^2\}}}, \dots \dots \dots (41)$$

where $\rho = |\zeta|$ is the distance of the point considered from the geometric centre, η its distance from the source line, and m the parameter.

The use of equation (40) leads to errors, if the field of φ in the w -plane is considerably curved. For practical work the case of importance is that one for which one of the two lines of constant φ , say the line $\varphi = \varphi_D$, is a straight line of constant u or v in the w -plane, corresponding to a boundary equipotential in the z -plane, and the other line of constant φ , that is, the line $\varphi = \varphi_A$, can be approximated by a circle. This is shown in Figures 12 (a) and (b) with the notations that correspond to the boundary equipotential in the z -plane being a line of constant v .

If coordinates ξ and η are introduced equalling u and v (Fig. 12 (a)) or $(-v)$ and u (Fig. 12 (b)) respectively, the field in the w -plane corresponding to Figure 12 (a) or 12 (b) can be assumed to be produced by a source and sink of equal intensities as was explained in Section III (a). If the intensity of the source and sink is C , then φ is given for Figure 12 (a) by the right-hand side of equation (13) multiplied by C , and for Figure 12 (b) by the right-hand side of

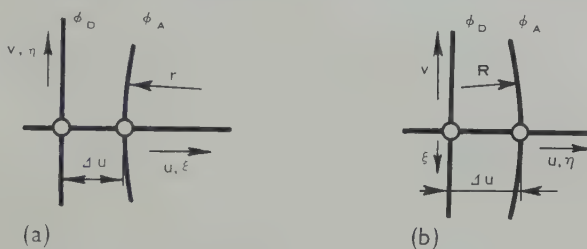


Fig. 12.—Equipotentials in the w -plane. (a) Case 1; (b) case 2.

equation (14) multiplied by C . A pair of these C -multiplied equations (13) or (14) written down for two pairs of values C, m and C', m' and equated for $\xi = \Delta u, \eta = 0$ (Fig. 12 (a)) or $\xi = 0, \eta = \Delta u$ (Fig. 12 (b)) gives a relation between C, m, C' , and m' for a fixed distance Δu and a fixed potential φ_A , but the equipotential lines φ_A are of different curvatures. If the C -multiplied right-hand side of equation (13) is differentiated with respect to ξ and ξ made equal to zero, and the C -multiplied right-hand side of equation (14) differentiated with respect to η and η made equal to zero, expressions for $\partial\varphi/\partial u$ along the equipotential φ_D result. The ratio c of $\partial\varphi/\partial u$ for a given value of m , to $(\partial\varphi/\partial u)'$ for $m' \rightarrow \infty$ can be computed and determines the increase or decrease of $\partial\varphi/\partial u$ compared with the case of a uniform field. If C is eliminated by using the relation between C, m, C' , and m' , C' cancels out and we obtain, after expressing m in terms of the radius of curvature r or R with the aid of equation (15) or (17) and an obvious relation eliminating b or B with the aid of Δu :

For Figure 12 (a)

$$c = \frac{\partial\varphi}{\partial u} \bigg/ \left(\frac{\partial\varphi}{\partial u} \right)' = \frac{m^2}{m^2 + \eta^2} \cdot c_0, \quad \dots\dots\dots (42)$$

$$c_0 = \frac{\Delta u/m}{\tanh^{-1} (\Delta u/m)}, \quad \dots\dots\dots (43)$$

$$\frac{\Delta u}{m} = \sqrt{\left(\frac{\Delta u}{2r + \Delta u} \right)}. \quad \dots\dots\dots (44)$$

For Figure 12 (b)

$$c = \frac{\partial \varphi}{\partial u} \left/ \left(\frac{\partial \varphi}{\partial u} \right)' \right. = \frac{m^2}{m^2 - \varepsilon^2} \cdot c_0, \quad \dots \dots \dots (45)$$

$$c_0 = \frac{\Delta u / m}{\tan^{-1} (\Delta u / m)}, \quad \dots \dots \dots (46)$$

$$\frac{\Delta u}{m} = \sqrt{\left(\frac{\Delta u}{2R - \Delta u} \right)}. \quad \dots \dots \dots (47)$$

c_0 is the ratio $(\partial \varphi / \partial u) / (\partial \varphi / \partial u)'$ at the line of symmetry of Figures 12 (a) and 12 (b) and is the value sought. $(\partial \varphi / \partial u)'$ is the value given by equation (40).



Fig. 13.—Example VIII (a). Field between two concentric circles. Coordinates of points:

A (3.9265, -1.6264)	F (6.8532, 2.8387)
B (4.2205, 0.50026)	G (9.2388, -3.8268)
C (3.9265, 1.6264)	H (9.9437, -1.0609)
D (5.5433, -2.2961)	I (9.2388, 3.8268)
E (6.6713, 0)	

(b) Squares with a Sharp Corner

For $\alpha > \pi$ the voltage gradient at the corner itself is infinite, but, as Cohn and Vogel (1953) have emphasized, its value at a short distance from the corner has significance in high voltage engineering. The voltage gradient at the corner for $\alpha < \pi$ is zero and its value in the neighbourhood of the corner is of little interest.

To compute the voltage gradient for a point at the distance ρ from the corner equations (38) and (39) are applied again, but $|dw/d\zeta|$ is given by

$$\left| \frac{dw}{d\zeta} \right| = \frac{1}{p} \cdot \left(\frac{t}{\rho} \right)^{(p-1)/p} \quad \dots\dots\dots (48)$$

Equation (48) is obtained from equation (32) by differentiation. For Δu and Δv the values t_1 and t respectively are taken (Fig. 9). The maximum gradient at the distance ρ from a corner of conducting material occurs for $\alpha > \pi$ on the bisector of the angle α and is directed from the corner. Since this is the only value of interest, no additional work regarding the directions of the lines $u=\text{constant}$ and $v=\text{constant}$ arises in this case.

VIII. EXAMPLES

Two simple examples without current loading ($\tau=0$) will demonstrate the power of the use of curvilinear nets. For the conductivity σ the value $\sigma=1$ is assumed. All numerical values given were computed with a computational accuracy to five digits. This accuracy is unnecessary, unless wanted for the purpose of comparison.

(a) Field between Two Concentric Circles

Figure 13 shows a portion of a sector of 45° aperture bounded by two concentric circle arcs ABC and GHI of radii 4.25 and 10 respectively as equipotentials and by two radii AG and CI as flow lines. This arrangement can easily be analysed by well-known formulae. We start arbitrarily at the point D at a distance 1.75 from the point A and, to simulate unfavourable conditions, select as the mesh boundary DE a circle arc of radius 20 with its centre on the line AD produced beyond the point D . The point E where the mesh boundary ends is the point of its intersection with the bisector of the sector. These assumptions determine a net of four meshes unequivocally. The computed characteristic data of the net are contained in Figure 13, the origin of the Cartesian frame used being the centre of the sector and the x -axis its bisector through E .

The values of u_1 , u_2 , Δu , v_1 , v_2 , Δv , $\Delta u/\Delta v$, and $\Delta v/\Delta u$ resulting for the individual meshes are shown in Table 2.

TABLE 2
VALUES OF u_1 , u_2 , Δu , v_1 , v_2 , Δv , $\Delta u/\Delta v$, AND $\Delta v/\Delta u$ FOR THE MESHES OF FIGURE 13

Mesh	<i>ABED</i>	<i>BCFE</i>	<i>DEHG</i>	<i>EFIH</i>
u_1	0.7923	0	0	1.6353
u_2	-0.1855	0.6956	0.7530	0.7296
Δu	0.9778	0.6956	0.7530	0.9057
v_1	0	0.7839	0.3526	0
v_2	1.2121	1.9463	-0.7626	1.2601
Δv	1.2121	1.1624	1.1152	1.2601
$\Delta u/\Delta v$	0.8067	0.5984	0.6752	0.7188
$\Delta v/\Delta u$	1.240	1.6720	1.4810	1.3912

If the scheme of Figure 4 is taken as the basis, and if the individual conductances are computed and all parallel conductances between two nodes are lumped, the network of Figure 14 results. The potentials φ_D , φ_E , and φ_F of the points D , E , and F were taken as unknowns and the potential of the node ($A-B-C$) was set equal to 1 and that of the node ($G-H-I$) set equal to zero. The results of the computations are given below and in the brackets are added the theoretically correct values and the per cent. deviations from them. Solving

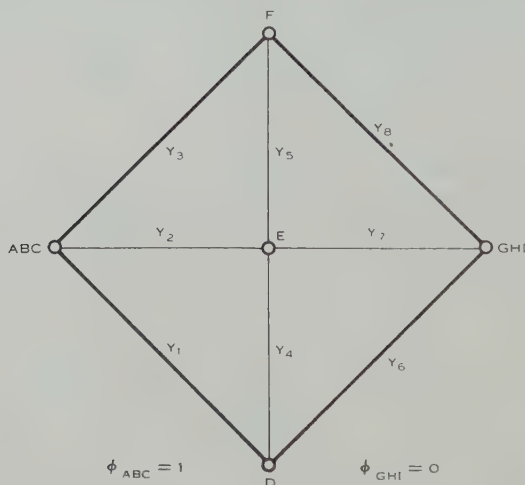


Fig. 14.—Values of the final conductances in examples VIII (a) and (b).

Conductance	Example VIII (a)	Example VIII (b)
Y_1	0.6200	0.6036
Y_2	0.9192	1.1036
Y_3	0.2992	0.5000
Y_4	0.6185	0.0550
Y_5	0.6478	0.5000
Y_6	0.3376	1.1004
Y_7	1.0332	1.6004
Y_8	0.6956	0.5000

three simultaneous linear equations expressing Kirchhoff's first rule leads to $\varphi_D=0.5781$ (0.5970, -3.2%), $\varphi_E=0.4707$ (0.4730, -0.57%), $\varphi_F=0.3678$ (0.3491, $+5.4\%$). It is somewhat misleading to consider the per cent. deviations from the actual values of the potentials, the value $1-\varphi_F=0.6322$ (0.6509, -2.9%) is more important, for instance, than φ_F . If the deviations are referred to the potential difference between the two outer electrodes the percentages are much smaller. The total conductance between the arcs ABC and GHI has the

value 0.9373 (0.9179 , $+2.1\%$). Hence the error in the computation of the capacity would be $+2.1\%$. With the aid of the formulae (38), (40), and (41) the approximate values of the gradients g_A at A and g_C at C were computed. The results are: $g_A=0.2696$ (0.2750 , -2.0%) and $g_C=0.3097$ (0.2750 , $+12.6\%$). Correction factors c_0 can be computed, for the node A with the



Fig. 15.—Example VIII (b). Field in a coaxial square cable. Coordinates of points :

A (0, 0)	F (1.0839, 6.0839)
B (0, 3.4672)	G (5, 0)
C (0, 5)	H (5, 4.4617)
D (3.2885, 0)	I (5, 10)
E (2.5, 3.9645)	

aid of equations (46) and (47), and for the node C with the aid of equations (43) and (44). The radius R required in equation (47) was ascertained by finding by linear interpolation in the w -plane the point of potential φ_D on the line $B'E'$. For the radius r required for equation (44) the point of potential

φ_E was similarly found on the line CF . We obtain $R=3.802$ and $r=1.3733$, $c_{0A}=1.0474$, $c_{0C}=0.9057$ and for the corrected values of the gradients $c_{0A} \cdot g_A=0.2824$ (0.2750 , $+2.7\%$) and $c_{0C} \cdot g_C=0.2805$ (0.2750 , $+2.0\%$).

Considering that the net consists of four meshes only and that the mesh boundaries deviate appreciably from the equipotentials and flow lines, all deviations from the correct values are surprisingly small.

(b) Field in a Square Coaxial Cable

A square conductor of side length 10 is surrounded by a square conducting sheath of side length 20 in an arrangement in which four axes of symmetry exist. This numerical problem has been dealt with by Woods (1953) and others. Figure 15 shows an eighth of the arrangement, AC being the considered portion of the inner equipotential, GI that of the outer equipotential, and AG and CI being flow lines. C is a sharp corner of 270° aperture and I one of 90° aperture. A is the origin of the selected Cartesian frame and AG its x -axis.

The condition that a net of four meshes be used determines the meshes unequivocally, the point E being the intersection of the (not drawn) bisectors of the angles ACI and CIG . Figure 15 contains the characteristic data of the net and Table 3 the values characteristic of the meshes.

TABLE 3
VALUES OF $u_1, u_2, \Delta u, v_1, v_2, \Delta v, \Delta u/\Delta v$ AND $\Delta v/\Delta u$ FOR THE MESHES OF FIGURE 15

Mesh	<i>ABED</i>	<i>BCFE</i>	<i>DEHG</i>	<i>EFIH</i>
$u_1 \dots \dots$	0	0	0	0
$u_2=\Delta u \dots \dots$	0.9894	1	0.5672	1
$v_1 \dots \dots$	0	0	0	0
$v_2=\Delta v \dots \dots$	0.8195	1	1.2483	1
$\Delta u/\Delta v \dots \dots$	1.2073	1	0.4544	1
$\Delta v/\Delta u \dots \dots$	0.8283	1	2.2008	1

The final resulting network of conductances is shown in Figure 14. The result of the analysis together with the figures ascertained by Woods (1953) or found from them by interpolation and the deviations from these figures that can be considered correct are given in the following :

$$\begin{aligned}\varphi_D &= 0.3562 \text{ (0.3326, } +7.0\%), \\ \varphi_E &= 0.4171 \text{ (0.4330, } -3.7\%), \\ \varphi_F &= 0.4724 \text{ (0.4950, } -4.6\%).\end{aligned}$$

The total conductance is 1.2957 (1.2791 , $+1.6\%$).

As in example VIII (a) the results are surprisingly good.

IX. REFERENCES

COHN, G. I., and VOGEL, F. J. (1953).—A solution of a common insulation problem. *Pwr. Apparatus and Systems* 6: 460.
MACNEAL, R. H. (1953).—An asymmetrical finite difference network. *Quart. Appl. Math.* 11: 295.

- MOTZ, H., and WORTHY, W. D. (1945).—Calculation of the magnetic field in dynamo-electric machines by Southwell's relaxation method. *J. Instn. Elect. Engrs.* II **92**: 522.
- SOUTHWELL, R. V. (1946).—"Relaxation Methods in Theoretical Physics." (Oxford Univ. Press.)
- TASNY-TSCHLIASSNY, L. (1949).—The triangulation of a two-dimensional continuum for the purpose of the approximate solution of second-order partial differential equations. *J. Appl. Phys.* **20**: 419.
- WOODS, L. C. (1953).—The relaxation treatment of singular points in Poisson's equation. *Quart. J. Mech.* **6**: 163, Example II.

SHARPENING OF OBSERVATIONAL DATA IN TWO DIMENSIONS

By E. J. BURR*

[Manuscript received August 20, 1954]

Summary

Distributions in two dimensions as measured are always blurred or smoothed by limitations in the observing technique. Recovery of the true distribution involves the solution of an integral equation of the form

$$g(x', y') = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y) f(x' - x, y' - y) dx dy,$$

where the functions g , h are known from observation.

In this paper the uniqueness and stability of the solution are discussed, and the validity and usefulness of several methods of solution are examined. A new technique for the application of polynomial solutions is presented. It is suggested that in some cases the use of Fourier transforms or Fourier-Bessel transforms may be practicable, and tables of selected transforms are supplied. Finally, conditions for convergence of certain sequences of approximate solutions are given. The effect of random errors in the observational data is taken into account wherever possible.

I. INTRODUCTION

Astronomical observations are sometimes concerned with distributions of surface intensity or statistical frequency distributions in two dimensions. Since it is not possible to measure a surface intensity or a frequency at a mathematical point, the limitations of the observational method always produce a smoothing or averaging effect on the true distribution, lowering and broadening the peaks; and reducing the amplitude of the fluctuations. The problem of recovering the true distribution from the smoothed distribution which we observe may be called "sharpening" of the observational data, and involves the solution of a linear integral equation.

The corresponding problem in one dimension has been frequently discussed, and it is recognized that the methods of solving this problem can in principle be extended to two dimensions, but this has rarely been done in practice. A solution in the form of a double series has been given by Coutrez (1949), but this does not appear to have been used. Kapteyn (1920) has found a special solution for use in the statistics of stellar total proper motions. Kreisel (1949) has described a method of partial sharpening which has been applied to a problem in gravity survey. Bolton and Westfold (1950) have solved a problem in radio astronomy with the aid of an iterative method described in one dimension by Burger and van Cittert (1932, 1933). A formal solution in terms of Fourier

* Commonwealth Observatory, Canberra, A.C.T.; present address: 23 Ruskin Street, Taringa, Brisbane, Qld.

transforms can be written down, but does not appear to have been used. Solutions by "trial and error" lead to uncertain results because of the indeterminacy caused by random errors of observation.

The object of this paper is to examine some of the methods of solving the integral equation in two dimensions, and devise means of applying them to numerical problems with the minimum of computational labour. The effect of random errors in the data is taken into account wherever possible.

II. THE INTEGRAL EQUATION

We denote the true distribution by $f(x, y)$ where (x, y) are rectangular Cartesian coordinates. The observed value at any point (x', y') is a weighted mean of the values of $f(x, y)$ in some neighbourhood of (x', y') , and we denote it by $g(x', y')$. Now suppose that, when the true distribution consists only of a point concentration of unit strength at (x_0, y_0) , the observed distribution is $h(x' - x_0, y' - y_0)$. Then the functions f, g, h are connected by the integral equation

$$g(x', y') = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x' - x, y' - y) f(x, y) dx dy,$$

or

$$g(x', y') = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y) f(x' - x, y' - y) dx dy. \quad \dots (2.1)$$

We shall call $h(x, y)$ the kernel, but in various contexts it may also be called the instrumental function, distribution of errors, or polar diagram. Equation (2.1) may be expressed in words by saying that $g(x, y)$ is the convolution, resultant, Faltung, obliteration, or smooth of $f(x, y)$ and $h(x, y)$. The process of recovering $f(x, y)$ from a knowledge of $g(x, y)$ and $h(x, y)$ may be called sharpening, restoration, or correction of the observed distribution $g(x, y)$.

We say that a function is L, or belongs to the class L, when its (Lebesgue) integral over the infinite plane converges absolutely. It is assumed throughout this paper that the kernel $h(x, y)$ is L. Then if $f(x, y)$ is bounded, and integrable in every finite region, the integral in (2.1) converges absolutely and uniformly, and $g(x', y')$ is bounded and continuous.

It is usual to *normalize* the kernel by multiplying it by a constant factor in order to make its total surface integral equal to unity. When this has been done, the total surface integrals of f and g are equal, provided that either of these functions is L. Loosely speaking, the two distributions f, g then differ in shape, but not in scale.

Equation (2.1) has the useful property that it is invariant under a homogeneous linear transformation of unit modulus. That is, if we write

$$\begin{aligned} x &= aX + bY, & ad - bc &= 1, \\ x' &= aX' + bY', & f(x, y) &\equiv f(aX + bY, cX + dY) \equiv \bar{f}(X, Y), \\ y &= cX + dY, & g(x, y) &\equiv \bar{g}(X, Y), \\ y' &= cX' + dY', & h(x, y) &\equiv \bar{h}(X, Y), \end{aligned}$$

then equation (2.1) reduces to

$$\bar{g}(X', Y') = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \bar{h}(X, Y) \bar{f}(X' - X, Y' - Y) dX dY.$$

This property enables particular solutions to be generalized, and certain problems to be simplified. For example, a problem in which one of the functions is a function of $(x^2/a^2 + y^2/b^2)$ only can be reduced to a problem in which this function has circular symmetry.

In astronomical applications, kernels of several different types occur in the following ways.

(i) In photoelectric photometry, an extended source of light may be scanned by a photocell behind a circular aperture of radius a in the focal plane of a telescope. f, g are the true and observed distributions of surface brightness, and the normalized kernel is

$$\left. \begin{aligned} h(x, y) &= 1/\pi a^2, & x^2 + y^2 < a^2, \\ &= 0, & x^2 + y^2 > a^2, \end{aligned} \right\} \dots\dots\dots (2.2)$$

provided that a is large compared with the effective resolving power of the telescope.

(ii) In photographic photometry, f is as in (i), and g is deduced from microphotometric tracings of a photograph of the source. The kernel is obtained from a stellar image by the same process. Many factors contribute to the smoothing process, the three most important being atmospheric turbulence, imperfect guiding of the telescope, and the finite size of the scanning "spot" in the microphotometer. As an example, measurements by de Vaucouleurs (1948) at the Cassegrain focus of a 32-in. reflector gave a kernel which is reasonably well represented by

$$\left. \begin{aligned} h(x, y) &= A e^{-\frac{1}{2}(x^2/a^2 + y^2/b^2)} \simeq A e^{-\frac{1}{2}r^2/c^2}, & r < 2.5, \\ &= 5 \cdot 2 A / r^4, & r > 2.5, \end{aligned} \right\} \dots\dots (2.3)$$

where $r^2 = x^2 + y^2$, the unit of measurement is 1 sec of arc, $a = 1.15$, $b = 1.35$, $c = 1.25$, and $A = 0.0904$. Equation (2.1) is strictly applicable only when stars in different parts of the field lead to the same kernel.

(iii) In radio astronomy, f, g are the true and observed distributions of surface intensity of an extended radio source. The kernel may be found by scanning a point source with the radio telescope, but this is often not practicable, and the theoretical diffraction pattern (polar diagram) of the telescope is then used. For example, for a circular aperture of diameter D , the diffraction pattern is

$$h(x, y) = \frac{A^2}{4\pi} \left[\frac{2}{Ar} J_1(Ar) \right]^2, \quad A = \frac{\pi D}{\lambda}, \quad \dots\dots\dots (2.4)$$

where λ is the wavelength of the radiation, $r^2 = x^2 + y^2$, and r is measured in radians. (This formula is only accurate for $r \ll 1$.) This kernel, in common with all diffraction patterns due to apertures of finite dimensions, has the important property that its Fourier transform vanishes outside a finite region.

(iv) In statistics, we take as a typical example the statistics of stellar proper motions. The number of stars in a given region whose true components of annual proper motion lie in the range $(x, x+dx; y, y+dy)$ is $f(x, y)dx dy$. The measures of (x, y) for individual stars are subject to random errors whose frequency distribution may be taken as

$$h_1(x, y) = (2\pi a^2)^{-1} e^{-\frac{1}{2}r^2/a^2}. \quad \dots\dots\dots (2.5)$$

The observed frequency distribution is smoothed by (2.5) and also by the necessity for taking samples over finite regions of the (x, y) plane. If $g(x', y')$ is taken as the number of stars with observed proper motions (x, y) in the range $|x-x'| \leq b, |y-y'| \leq c$, then $g(x', y')$ is given by (2.1) with a kernel obtained from the convolution of (2.5) with

$$h_2(x, y) = \begin{cases} (4bc)^{-1}, & |x| < b, |y| < c \\ 0, & \text{elsewhere.} \end{cases} \quad \dots\dots\dots (2.6)$$

We may also mention a problem in gravity survey (Kreisel 1949) in which the kernel is the normal gravitational field over a horizontal plane due to unit mass at depth a below the origin, which is

$$h(x, y) = Ga(a^2 + r^2)^{-3/2}. \quad \dots\dots\dots (2.7)$$

In all these applications the distributions are, strictly speaking, over a sphere, not a plane. But in nearly every case the angular diameter of the region with which we have to deal is so small that the region may be regarded as plane without sensible error.

III. THE FORMAL SOLUTION

Any function $f(x, y)$ of L has a double Fourier transform (Bochner 1932)

$$F(u, v) = (2\pi)^{-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{i(ux+vy)} dx dy, \quad \dots\dots\dots (3.1)$$

which is bounded and continuous and tends to zero as $u^2 + v^2 \rightarrow \infty$. The inverse relation

$$f(x, y) = (2\pi)^{-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v) e^{-i(ux+vy)} du dv \quad \dots\dots (3.2)$$

holds wherever $f(x, y)$ is continuous; provided that, if the integral in (3.2) does not converge absolutely, we interpret it as the limit as $p, q \rightarrow \infty$ of the integral over $|u| < p, |v| < q$. If two functions belonging to L have the same Fourier transform, they are equal almost everywhere, so that, if they represent distributions of a physical quantity, these distributions are identical.

Now let $f(x, y), h(x, y)$ be L ; then $g(x, y)$ of equation (2.1) is also L , and, if we denote Fourier transforms by capital letters, the Faltung theorem (Bochner 1932, §43.6) gives

$$G(u, v) = 2\pi H(u, v) F(u, v). \quad \dots\dots\dots (3.3)$$

This leads at once to a formal solution in the form (3.2), in which

$$F(u, v) = \frac{G(u, v)}{2\pi H(u, v)}. \quad \dots\dots\dots (3.4)$$

Now suppose that g, h are known from observation, and belong to L . Then we can (in principle at least) compute their Fourier transforms and form the quotient (3.4). If $|H(u, v)|$ is everywhere positive, as is the case with kernels such as (2.5) and (2.7), then (3.4) is defined everywhere, and, if this function has a Fourier transform belonging to L , the solution exists and is unique. But if $H(u, v)$ vanishes at certain real points, then $G(u, v)$ ought also to vanish at these points and (3.4) is undefined at such points. (Note that we are concerned only with real values of (u, v) , although F, G, H are in general complex functions of these variables.) Denote by R the region or point set in which $H(u, v) = 0$. We may arbitrarily set $F(u, v) = 0$ in R , and, if $F(u, v)$ then has a Fourier transform, this transform is a possible solution of the integral equation, which Bracewell and Roberts (1954) call the *principal solution*. Other solutions may be obtained by adding to the principal solution functions of the form

$$\iint_R \varphi(u, v) e^{-i(ux+vy)} du dv, \dots\dots\dots (3.5)$$

where $\varphi(u, v)$ is an arbitrary function. If R is a set of zero measure, as is the case with kernels such as (2.2) and (2.6), then (3.5) vanishes unless $\varphi(u, v)$ is an improper function, and in many applications the indeterminacy is then somewhat trivial. But if R has positive measure the indeterminacy is much more serious. This occurs with kernels like (2.4) which are of special importance in radio astronomy, and the problems raised by this indeterminacy have been fully discussed in the one-dimensional case by Bracewell and Roberts (1954). Their discussion can easily be generalized to two dimensions.

Fourier transforms may be expressed in polar coordinates by setting

$$\begin{aligned} x &= r \cos \theta, & u &= s \cos \varphi, \\ y &= r \sin \theta, & v &= s \sin \varphi. \end{aligned}$$

Then (3.1), (3.2) take the form

$$\left. \begin{aligned} F(s, \varphi) &= (2\pi)^{-1} \int_0^\infty \int_{-\pi}^\pi f(r, \theta) e^{isr \cos(\theta - \varphi)} r dr d\theta, \\ f(r, \theta) &= (2\pi)^{-1} \int_0^\infty \int_{-\pi}^\pi F(s, \varphi) e^{-isr \cos(\theta - \varphi)} s ds d\varphi. \end{aligned} \right\} \dots\dots\dots (3.6)$$

In particular, when $f(r, \theta)$ is independent of θ , it is clear that $F(s, \varphi)$ is independent of φ , and these functions may be written as $f(r)$, $F(s)$. Using Parseval's integral for the Bessel function (Watson 1944), the relations (3.6) then become

$$F(s) = \int_0^\infty f(r) J_0(sr) r dr, \dots\dots\dots (3.7)$$

$$f(r) = \int_0^\infty F(s) J_0(sr) s ds. \dots\dots\dots (3.8)$$

The functions $f(r)$, $F(s)$ are called Fourier-Bessel transforms of one another.

If the functions f, g, h of equation (2.1) all have circular symmetry, they may be written $f(r), g(r), h(r)$, and (3.3) becomes

$$G(s) = 2\pi H(s)F(s),$$

so that the formal solution has the form (3.8), in which

$$F(s) = \frac{G(s)}{2\pi H(s)}.$$

IV. INSTABILITY OF THE SOLUTION

If g, h are specified exactly, as in a theoretical problem, there may be a unique and exact solution. But if the values of $g(x, y)$ are uncertain due to observational errors, however small these errors may be, the solution is always completely indeterminate even with the restriction that $f(x, y) \rightarrow 0$ as $x^2 + y^2 \rightarrow \infty$. For let any solution be given an (additive) increment of the form

$$\Delta f(x, y) = A e^{-i(ux+vy)} w(x, y),$$

where

$$w(x, y) = (1 - |x|/a)(1 - |y|/b), \quad |x| < a, \quad |y| < b, \\ = 0 \quad \text{elsewhere,}$$

and A, a, b are arbitrary. Then it may be shown that the corresponding increment of g can be made arbitrarily small by taking $u^2 + v^2$ sufficiently large. This property may be expressed by saying that $f(x, y)$ is not continuously determined by $g(x, y)$. (Compare Kreisel 1949.)

In other words, oscillating increments of arbitrarily large amplitude in $f(x, y)$ correspond to unobservably small increments in $g(x, y)$ provided that the wavelength of the oscillations is sufficiently small. Such oscillations appear when the computer tries to determine $f(x, y)$ with greater precision than is warranted by the accuracy of the observations.

The solution must therefore be stabilized by excluding functions which have oscillations of large amplitude and short wavelength. This means that complete sharpening can never be achieved except in theoretical problems. Following a suggestion by Fellgett and Schmeidler (1952), if we are given the autocorrelation function of the errors in $g(x, y)$, we can use the Wiener-Kolmogoroff smoothing theory (Bode and Shannon 1950) to determine a solution giving the best possible compromise between errors due to magnification of the errors in $g(x, y)$ and errors due to incomplete sharpening. But the difficulty of measuring the autocorrelation function would seem to make this method of little practical value. It is therefore necessary to use more arbitrary methods of stabilization. Unfortunately this means that the solution adopted is to some extent subjective, that is, depending on the computer's judgment. But, if the statement of the solution is accompanied by a statement as to which method of sharpening was used, the result so expressed is purely objective, and is therefore to be preferred to a solution reached by "trial and error".

When the highest possible accuracy is required, it appears that the best procedure is to find a sequence of approximate solutions of (theoretically)

increasing accuracy, taking the observational data at their face value ; and then to terminate the sequence just before physically improbable undulations begin to appear.

V. POLYNOMIAL SOLUTIONS

If $f(x, y)$ is the polynomial

$$f(x, y) = \sum_{m=0}^A \sum_{n=0}^B f_{mn} x^m y^n / m! n!, \quad \dots \dots \dots (5.1)$$

then the integral in (2.1) converges if $h(x, y)$ possesses moments of all requisite orders. We define these moments by

$$M_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{x^p y^q}{p! q!} h(x, y) dx dy. \quad \dots \dots \dots (5.2)$$

The integral in (2.1) can then be evaluated either by term-by-term binomial expansion of $f(x' - x, y' - y)$, or by Taylor expansion of this function in powers of x and y . Thus $g(x', y')$ may be expressed in either of the two forms

$$g(x', y') = \sum_{m=0}^A \sum_{n=0}^B f_{mn} \sum_{p=0}^m \sum_{q=0}^n (-1)^{m+n-p-q} M_{m-p, n-q} x'^p y'^q / p! q! \quad \dots (5.3)$$

$$= \sum_{m=0}^A \sum_{n=0}^B (-1)^{m+n} M_{mn} \frac{\partial^{m+n}}{\partial x'^m \partial y'^n} f(x', y'). \quad \dots \dots \dots (5.4)$$

Similarly, if $g(x, y)$ is the polynomial

$$g(x, y) = \sum_{m=0}^A \sum_{n=0}^B g_{mn} x^m y^n / m! n!, \quad \dots \dots \dots (5.5)$$

it may be verified by substitution that (2.1) is satisfied by

$$f(x, y) = \sum_{m=0}^A \sum_{n=0}^B g_{mn} \sum_{p=0}^m \sum_{q=0}^n (-1)^{m+n-p-q} \mu_{m-p, n-q} x^p y^q / p! q! \quad \dots (5.6)$$

$$= \sum_{m=0}^A \sum_{n=0}^B (-1)^{m+n} \mu_{mn} \frac{\partial^{m+n}}{\partial x^m \partial y^n} g(x, y), \quad \dots \dots \dots (5.7)$$

where the inverse moments μ_{pq} are defined by

$$\left. \begin{aligned} \mu_{00} M_{00} &= 1, \\ \sum_{p=0}^m \sum_{q=0}^n \mu_{m-p, n-q} M_{pq} &= 0, \quad (m+n > 0). \end{aligned} \right\} \quad \dots \dots \dots (5.8)$$

Thus the solution (5.6) is valid provided that $M_{00} \neq 0$, and it is easy to show that it is the only polynomial solution. A result equivalent to (5.7) has been obtained by Coutrez (1949), and the one-dimensional form of (5.6) has been given by van de Hulst (1946). (The moments as defined by these authors correspond to $p!q!M_{pq}$ in our notation.)

If the integrals (5.2) converge absolutely for all non-negative integers p, q , then the integral defining $2\pi H(u, v)$ may be differentiated any number of times

under the integral sign. We then find that $2\pi H(u, r)$ and all its derivatives are bounded and continuous for all (u, r) , so that the Maclaurin expansion

$$2\pi H(u, r) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} M_{mn} (iu)^m (ir)^n \dots\dots\dots (5.9)$$

converges. From (5.8) and (5.9),

$$1/2\pi H(u, v) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \mu_{mn} (iu)^m (iv)^n \dots\dots\dots (5.10)$$

The relations (5.9) and (5.10) facilitate computation of the moments and inverse moments when $H(u, r)$ is expressible in terms of elementary functions. Note that, for a normalized kernel, $2\pi H(0, 0) = M_{00} = \mu_{00} = 1$.

The results (5.3)–(5.7) remain true when the polynomials are extended to infinite series provided that the series for $h(x, y)f(x' - x, y' - y)$ converges for all (x, y) and may be integrated term by term over the infinite (x, y) plane. However, we are here concerned primarily with distributions given by observational data, and the need for infinite series does not arise.

The results (5.4) and (5.7) may be formally applied to functions which are not polynomials. In (5.4) it is usually possible to write down a remainder term setting an upper limit to the error caused by neglecting higher derivatives, but in (5.7) there is no simple way of estimating this error. However, if $f(x, y)$ and all its moments exist, it may be shown that the moments of the right-hand member of (5.7) are the same as the moments of the true solution $f(x, y)$ up to and including moments of order (A, B) , so that in this narrow sense every term of (5.7) effects an improvement in the approximation to $f(x, y)$ as a whole. (Compare Eddington's treatment (1940) of the one-dimensional problem with Gaussian kernel.) But it does not necessarily follow that the series converges at any particular point (x, y) .

VI. POLYNOMIAL SOLUTIONS WITH CIRCULAR SYMMETRY

If f, g, h are functions of r only, where $r^2 = x^2 + y^2$, we may write them as $f(r), g(r), h(r)$ and the integral equation (2.1) becomes

$$g(r') = \int_0^{\infty} \int_{-\pi}^{\pi} h(r) f[\sqrt{(r'^2 + r^2 - 2rr' \cos \theta)}] r dr d\theta. \dots (6.1)$$

If $f(r)$ is a polynomial of degree A in r^2 ,

$$f(r) = \sum_{n=0}^A f_n r^{2n} / n!^2, \dots\dots\dots (6.2)$$

then the integral in (6.1) converges if $h(r)$ possesses *circular moments* of orders up to A . We define these circular moments by

$$M_p = 2\pi \int_0^{\infty} \frac{r^{2p}}{p!^2} h(r) r dr. \dots\dots\dots (6.3)$$

Observe that, by multiplying the binomial expansions of $(r' - re^{i\theta})^n/n!$ and $(r' - re^{-i\theta})^n/n!$, we obtain

$$\frac{(r'^2 + r^2 - 2rr' \cos \theta)^n}{n!^2} = \sum_{p=0}^n \sum_{q=0}^n \frac{r'^{p+q}}{p!q!} \frac{(-r)^{2n-p-q}}{(n-p)!(n-q)!} \cos (q-p)\theta,$$

so that

$$\int_{-\pi}^{\pi} \frac{(r'^2 + r^2 - 2rr' \cos \theta)^n}{n!^2} d\theta = 2\pi \sum_{p=0}^n \frac{r'^{2p}}{p!^2} \frac{r^{2n-2p}}{(n-p)!^2}. \quad \dots\dots\dots (6.4)$$

Then on substituting from (6.2) in (6.1) and carrying out the integrations with the aid of (6.4) and (6.3), we obtain

$$g(r') = \sum_{n=0}^A f_n \sum_{p=0}^n M_{n-p} r'^{2p}/p!^2. \quad \dots\dots\dots (6.5)$$

Similarly, if $g(r)$ is the polynomial

$$g(r) = \sum_{n=0}^A g_n r^{2n}/n!^2, \quad \dots\dots\dots (6.6)$$

it may be verified by substitution that (6.1) is satisfied by

$$f(r) = \sum_{n=0}^A g_n \sum_{p=0}^n \mu_{n-p} r^{2p}/p!^2, \quad \dots\dots\dots (6.7)$$

where the *inverse circular moments* μ_p are defined by

$$\left. \begin{aligned} \mu_0 M_0 &= 1, \\ \sum_{p=0}^n \mu_{n-p} M_p &= 0, \quad n > 0. \end{aligned} \right\} \quad \dots\dots\dots (6.8)$$

Thus the solution (6.7) is valid provided that $M_0 \neq 0$, and it is easy to show that it is the only polynomial solution.

We may express (6.5) and (6.7) in terms of the derivatives of f, g , by using the differential operator D defined by

$$D\varphi(r) = \frac{d}{d(r^2)} \left[r^2 \frac{d\varphi}{d(r^2)} \right] = \frac{1}{4r} \frac{d}{dr} \left(r \frac{d\varphi}{dr} \right) = \frac{1}{4} \nabla^2 \varphi.$$

Equations (6.5) and (6.7) then become

$$g(r') = \sum_{n=0}^A M_n D^n f(r'), \quad \dots\dots\dots (6.9)$$

$$f(r) = \sum_{n=0}^A \mu_n D^n g(r). \quad \dots\dots\dots (6.10)$$

If the integrals (6.3) converge absolutely for all p , then the integral defining $2\pi H(s)$ may be differentiated any number of times under the integral sign, and we find that

$$2\pi H(s) = \sum_{n=0}^{\infty} M_n (-s^2/4)^n, \quad \dots\dots\dots (6.11)$$

$$1/2\pi H(s) = \sum_{n=0}^{\infty} \mu_n (-s^2/4)^n. \quad \dots\dots\dots (6.12)$$

In this case the kernel also possesses all Cartesian moments M_{mn} , and these can be expressed in terms of the circular moments M_n . For if in (6.11) we write $s^{2n} = (u^2 + v^2)^n$, expand each term binomially, and identify the result with (5.9), we obtain

$$\left. \begin{aligned} M_{2m+1, n} &= M_{m, 2n+1} = 0, \\ M_{2m, 2n} &= \frac{(m+n)! M_{m+n}}{m! n! 4^{m+n}}. \end{aligned} \right\} \dots\dots\dots (6.13)$$

Similarly from (6.12) and (5.10) we obtain

$$\left. \begin{aligned} \mu_{2m+1, n} &= \mu_{m, 2n+1} = 0, \\ \mu_{2m, 2n} &= \frac{(m+n)! \mu_{m+n}}{m! n! 4^{m+n}}. \end{aligned} \right\} \dots\dots\dots (6.14)$$

VII. APPLICATION OF POLYNOMIAL SOLUTIONS

Let $h(x, y)$ be given and possess moments of at least the first few orders, which are computed from (5.2). The inverse moments are then found from (5.8). Let $g(x, y)$ be given by a table of its values at an array of points (pa, qb) where p, q are any integers and a, b are constants. To compute an approximate value of f at any one of these points, we may temporarily regard this point as the origin of coordinates without loss of generality, so that, from (5.6),

$$f(0, 0) = \Sigma \Sigma (-1)^{m+n} g_{mn} \mu_{mn}, \dots\dots\dots (7.1)$$

where the coefficients g_{mn} are determined so that the polynomial $\Sigma \Sigma g_{mn} x^m y^n / m! n!$ coincides with the tabulated g -values at a set of points (pa, qb) in the neighbourhood of the (temporary) origin. The number of points in this set equals the number of coefficients g_{mn} , and it is clear that this point set should extend over a region at least as large as the region in which $h(x, y)$ differs appreciably from zero. Then it is found that (7.1) is a homogeneous linear function of the g -values at these points, whose coefficients depend only on the inverse moments and on the geometry of the set of points. For a particular kernel and point set, these coefficients constitute a set of weights which need be numerically evaluated only once and can be used repeatedly to correct different observed distributions or different parts of the same observed distribution.

To illustrate the method, we consider in more detail kernels which have two orthogonal axes of symmetry. Taking these as coordinate axes, we find that $M_{mn} = 0$, $\mu_{mn} = 0$ whenever m or n is odd. Then (7.1) becomes

$$f(0, 0) = \Sigma \Sigma g_{2m, 2n} \mu_{2m, 2n}.$$

Evidently those terms of the polynomial which contain an odd power of x or of y contribute nothing to $f(0, 0)$, so that the correction of $g(x, y)$ is equivalent to the correction of the function

$$\bar{g}(x, y) = \frac{1}{4} [g(x, y) + g(-x, y) + g(x, -y) + g(-x, -y)].$$

Now $\bar{g}(x, y)$ is fully determined by its values in the first quadrant of the (x, y) plane, so that if the point set is symmetrical with respect to both axes, then the

number of points within or bordering the first quadrant must be equal to the number of terms in the polynomial for $\bar{g}(x, y)$. We take this polynomial to be complete and of degree N in (x^2, y^2) , that is, it contains all terms of the form $g_{2m, 2n} x^{2m} y^{2n}$ for $0 \leq m+n \leq N$. The number of terms in this polynomial is $\frac{1}{2}(N+1)(N+2)$.

We take for our point set all points (pa, qb) for which $|p| + |q| \leq N$; that is, the points within or bordering a rhombus whose vertices are $(\pm Na, 0)$, $(0, \pm Nb)$. There are $2N^2 + 2N + 1$ such points, of which $\frac{1}{2}(N+1)(N+2)$ are within or bordering the first quadrant, as required. By generalizing Lagrange's interpolation formula, the polynomial for $\bar{g}(x, y)$ may be written

$$\bar{g}(x, y) = \sum_{p=0}^N \sum_{q=0}^{N-p} g(pa, qb) \varphi_{pq}(x, y) / \varphi_{pq}(pa, qb), \dots \quad (7.2)$$

where

$$\varphi_{pq}(x, y) = (x^2 - p^2 a^2)^{-1} (y^2 - q^2 b^2)^{-1} \prod_{m=0}^N \prod_{n=0}^{N-m} (x^2 - m^2 a^2) (y^2 - n^2 b^2).$$

To complete the computation of the weights for a particular N , we have now only to expand (7.2) in polynomial form and then replace $x^{2m} y^{2n} / (2m)!(2n)!$ by $\mu_{2m, 2n}$. Writing the inverse moments in the dimensionless form

$$m_{2p, 2q} = \mu_{2p, 2q} / a^{2p} b^{2q},$$

the result for $N=1$, involving five points, is

$$f(0, 0) = g(0, 0)(m_{00} - 2m_{20} - 2m_{02}) + \bar{g}(a, 0)2m_{20} + \bar{g}(0, b)2m_{02}, \dots \quad (7.3)$$

and for $N=2$, involving 13 points,

$$\begin{aligned} 6f(0, 0) = & g(0, 0)(6m_{00} - 15m_{20} + 36m_{40} + 24m_{22} + 36m_{04}) \\ & + \bar{g}(a, 0)(16m_{20} - 48m_{40} - 24m_{22}) + \bar{g}(0, b)(16m_{02} - 48m_{04} - 24m_{22}) \\ & + \bar{g}(2a, 0)(-m_{20} + 12m_{40}) + \bar{g}(0, 2b)(-m_{02} + 12m_{04}) + \bar{g}(a, b)24m_{22} \\ & \dots \quad (7.4) \end{aligned}$$

The formula for $N=3$, involving 25 points, is so unwieldy that it is not likely to be of much practical value.

Expressions for $g(0, 0)$ in terms of the values of f at a rhomboidal array of points can also be derived in the same way. They may be obtained immediately from (7.3) and (7.4) by interchanging the symbols f and g , and substituting $M_{2p, 2q}$ for $\mu_{2p, 2q}$.

When the kernel has circular symmetry and $a=b$, the formulae may be simplified by expressing the moments in terms of the circular moments according to (6.14). Writing $m_p = \mu_p / a^{2p}$, the formulae (7.3) and (7.4) then become

$$f(0, 0) = g(0, 0)(m_0 - m_1) + \bar{g}(a, 0)\frac{1}{2}m_1 + \bar{g}(0, a)\frac{1}{2}m_1, \dots \quad (7.5)$$

$$\begin{aligned} 24f(0, 0) = & g(0, 0)(24m_0 - 30m_1 + 30m_2) + [\bar{g}(a, 0) + \bar{g}(0, a)](16m_1 - 24m_2) \\ & + [\bar{g}(2a, 0) + \bar{g}(0, 2a)](-m_1 + 3m_2) + \bar{g}(a, a)12m_2. \dots \quad (7.6) \end{aligned}$$

The formulae (7.3), (7.5) are exact if $g(x, y)$ is a polynomial of degree 3 in (x, y) and (7.4), (7.6) are exact if $g(x, y)$ is a polynomial of degree 5 in (x, y) . If the observed data $g(pa, qb)$ contain random (additive) errors of mean zero and mean square ε^2 , then the weighted sums will contain errors of mean zero and mean square $\lambda^2\varepsilon^2$, where λ^2 is the sum of the squares of the $(2N^2+2N+1)$ weights in the expression for $f(0, 0)$. If we increase N without limit, keeping aN, bN constant, we find that λ^2 increases without limit. This is a manifestation of the instability discussed in Section IV. In practice we implicitly exclude rapidly undulating functions by approximating to f, g by polynomials of low degree, that is, by keeping N small. Thus we obtain stable solutions, but the sharpening is incomplete when $g(x, y)$ is not a polynomial, or is a polynomial whose degree exceeds $2N+1$ in (x, y) .

For the Gaussian kernel

$$h(r) = (2\pi c^2)^{-1} e^{-\frac{1}{2}r^2/c^2}, \quad \dots\dots\dots (7.7)$$

the circular moments are $M_n = (2c^2)^n/n!$, $\mu_n = (-2c^2)^n/n!$, and weights for the correction $f-g$, computed from (7.5) and (7.6), are given in Table 1. The meaning of the weights for $N=1$, written in full, is

$$f(x, y) - g(x, y) = (c^2/2a^2)[4g(x, y) - g(x+a, y) - g(x-a, y) - g(x, y+a) - g(x, y-a)],$$

from which the meaning of the other weights should be clear.

For the kernel representing a circular aperture

$$h(r) = \begin{cases} 1/\pi c^2 & (r < c) \\ = 0 & (r > c). \end{cases} \quad \dots\dots\dots (7.8)$$

the circular moments are $M_n = c^{2n}/n!(n+1)!$, $\mu_0 = 1$, $\mu_1 = -\frac{1}{2}c^2$, $\mu_2 = c^4/6$, $\mu_3 = -7c^6/144$, $\mu_4 = 13c^8/960$, . . . , and sets of weights for the correction $f-g$ are given in the lower half of Table 1.

When f, g, h all have circular symmetry, a similar method may be used to find formulae involving only g -values at suitably spaced points along a diameter. We choose $2N+1$ equally spaced points extending from $r=(n-N)a$ to $r=(n+N)a$, $n \geq N$, and fit a polynomial of degree $2N$ in r^2 at these points. Lagrange's interpolation formula then gives

$$g(r) = \sum_{p=-N}^N g(na+pa) \varphi_p(r) / \varphi_p(na+pa), \quad \dots\dots\dots (7.9)$$

where

$$\varphi_p(r) = [r^2 - (n+p)^2 a^2]^{-1} \prod_{j=-N}^N [r^2 - (n+j)^2 a^2].$$

We now expand (7.9) in polynomial form and then replace $r^{2m}/m!^2$ by $\sum_{i=0}^m \mu_{m-i} r^{2i}/i!^2$ to obtain $f(r)$. To obtain $f(na)$, $n \geq N$, we set $r=na$ in the result. The expression for $f(r)$ derived from the points $0, a, 2a, \dots, 2Na$ may be taken as the best approximation to $f(r)$ for all $r \leq Na$. Thus we obtain sets of weights for $f(0), f(a), f(2a), \dots$. Unfortunately these sets are all different, because the

TABLE 1
WEIGHTS FOR COMPUTING $f(x, y) - g(x, y)$

Weights with the kernel (7.7)									
$N=1$			$N=2, a=2c$					$N=2, a=c\sqrt{2}$	
			7					5	
-1			6 -88 6					6 -56 6	
-1	4	-1	7	-88	300	-88	7	5	-56 180 -56 5
-1			6 -88 6					6 -56 6	
			7					5	
Multiply by $c^2/2a^2$ If $a=2c, \lambda=1.52$			Multiply by $1/384$ $\lambda=1.84$					Multiply by $1/96$ $\lambda=3.11$	

Weights with the kernel (7.8)									
$N=1$			$N=2, a=c$					$N=2, a=c/\sqrt{2}$	
			1					3	
-1			1 -12 1					4 -32 4	
-1	4	-1	1	-12	40	-12	1	3	-32 100 -32 3
-1			1 -12 1					4 -32 4	
			1					3	
Multiply by $c^2/8a^2$ If $a=c, \lambda=1.52$			Multiply by $1/48$ $\lambda=1.90$					Multiply by $1/48$ $\lambda=3.29$	

origin can no longer be chosen arbitrarily. For $N=1$, involving three points, the general formula is

$$\begin{aligned}
 4n(4n^2-1)[f(r)-m_0g(r)] &= (2n+1)g(\overline{n-1}a)[4m_2-m_1(2n^2+2n+1-4r^2/a^2)] \\
 &\quad -4n g(na)[4m_2-m_1(2n^2+2-4r^2/a^2)] \\
 &\quad + (2n-1)g(\overline{n+1}a)[4m_2-m_1(2n^2-2n+1-4r^2/a^2)],
 \end{aligned}$$

where $m_p = \mu_p/a^{2p}$. For the particular kernel (7.8), with $a=c$, the first few formulae for $N=1$ are:

$$\begin{aligned}
 f(0)-g(0) &= +0.792g(0) -0.889g(a) +0.097g(2a), \\
 f(a)-g(a) &= +0.292g(0) -0.222g(a) -0.069g(2a), \\
 f(2a)-g(2a) &= -0.035g(a) +0.156g(2a) -0.121g(3a), \\
 f(3a)-g(3a) &= -0.081g(2a) +0.210g(3a) -0.129g(4a), \\
 f(4a)-g(4a) &= -0.097g(3a) +0.228g(4a) -0.131g(5a), \\
 f(5a)-g(5a) &= -0.105g(4a) +0.236g(5a) -0.131g(6a),
 \end{aligned}$$

where the coefficients are correct to three decimal places. For the same kernel with $a=c$ and $N=2$, we find :

$$\begin{aligned} f(0)-g(0) &= +1.132g(0) - 1.414g(a) + 0.331g(2a) - 0.053g(3a) + 0.004g(4a), \\ f(a)-g(a) &= +0.086g(0) + 0.095g(a) - 0.210g(2a) + 0.031g(3a) - 0.002g(4a), \\ f(2a)-g(2a) &= -0.576g(0) + 0.808g(a) - 0.147g(2a) - 0.091g(3a) + 0.006g(4a), \\ f(3a)-g(3a) &= -0.084g(a) + 0.058g(2a) + 0.169g(3a) - 0.155g(4a) + 0.013g(5a). \end{aligned}$$

For comparison, we give three formulae for correction of one-dimensional distributions when the kernel represents a slit of width $2c$, that is,

$$\begin{aligned} h(x) &= 1/2c & (|x| < c) \\ &= 0 & (|x| > c). \end{aligned}$$

The three-point formula is then

$$f(x)-g(x) = (c^2/6a^2)[-g(x-a) + 2g(x) - g(x+a)].$$

When $a=c$, the five-point formula is

$$30[f(x)-g(x)] = g(x-2a) - 9g(x-a) + 16g(x) - 9g(x+a) + g(x+2a),$$

and when $a=c/\sqrt{2}$,

$$180[f(x)-g(x)] = 19g(x-2a) - 136g(x-a) + 234g(x) - 136g(x+a) + 19g(x+2a).$$

It should be stressed that the formulae of this section usually give accurate results only when the corrections are small. This is partly because the observational data cannot usually be well represented by polynomials of low degree, and partly because we have characterized the kernel by only a few parameters.

VIII. THE USE OF FOURIER TRANSFORMS

The formal solution by Fourier transforms, given in Section III, is hardly suitable for direct numerical computations. But when $g(x, y)$ and $h(x, y)$ can be closely approximated by elementary functions whose transforms are known, this form of solution may be useful. Also, in radio astronomy, measurements with a variable-spacing interferometer give $|F(u, v)|^2$ directly over a finite region, from which hypotheses concerning $f(x, y)$ may be tested.

Some Fourier transforms likely to prove useful for these purposes are listed in Table 2. The two columns of this table may be interchanged by interchanging (u, v) with (x, y) and writing $-i$ for i . In lines (1) and (5), F_1, F_2 are the Fourier transforms of f_1, f_2 . Lines (1)–(7) give rules for extending known transforms. Lines (8)–(11) enable us to find the transforms of peak functions and of periodic functions. All the results in Table 2 are well known or have obvious proofs, possibly excepting line (14). By using the relations

$$(d/dx)H_n(x) = nH_{n-1}(x) = xH_n(x) - H_{n+1}(x)$$

we can prove by induction that the (simple) Fourier transform of $H_n(a_1x)e^{-\frac{1}{2}x^2}$ is $(i\alpha)^n H_n(a_1\alpha/x)e^{-\frac{1}{2}u^2}$, where $\alpha^2 = a_1^2 - 1$; from which line (14) follows. We give here the first few of the Hermite polynomials :

$$\begin{aligned} H_0(x) &= 1, & H_1(x) &= x, & H_2(x) &= x^2 - 1, & H_3(x) &= x^3 - 3x, \\ H_4(x) &= x^4 - 6x^2 + 3, & H_5(x) &= x^5 - 10x^3 + 15x, \\ H_6(x) &= x^6 - 15x^4 + 45x^2 - 15, & H_n(x) &= n! \sum_p (-1)^p x^{n-2p} / 2^p p! (n-2p)! \end{aligned}$$

For brevity, we write

$$\varphi(a^2, b^2) = (2\pi ab)^{-1} e^{-\frac{1}{2}(x^2/a^2 + y^2/b^2)}. \quad \dots\dots\dots (8.1)$$

Then, if $h(x, y) = \varphi(a^2, b^2)$, the solution of the integral equation (2.1) can be written down whenever $g(x, y)$ is expressible as one of the functions on lines (12)–(15) of Table 2, or as a sum of such functions. For example, if

$$g(x, y) = \Sigma \Sigma c_{mn} \varphi(a_m^2, b_n^2), \quad \dots\dots\dots (8.2)$$

TABLE 2
FOURIER TRANSFORMS

The functions in the right-hand column are Fourier transforms of those in the left-hand column.
The Hermite polynomials are defined by $(d/dx)^n e^{-\frac{1}{2}x^2} = (-1)^n H_n(x) e^{-\frac{1}{2}x^2}$

$(2\pi)^{-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-i(ux+vy)} du dv$	$(2\pi)^{-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v) e^{i(ux+vy)} dx dy$
(1) $af_1(x, y) + bf_2(x, y)$	$aF_1(u, v) + bF_2(u, v)$
(2) $f(x-a, y-b)$	$F(u, v) e^{i(au+bv)}$
(3) $f(x/a, y/b)$	$abF(au, bv)$
(4) $f(ax+by, cx+dy), ad-bc=\Delta$	$\Delta^{-1} F\{(du-cv)/\Delta, (-bu+av)/\Delta\}$
(5) $2\pi f_1(x, y) f_2(x, y)$	$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_1(U, V) F_2(u-U, v-V) dU dV$
(6) $(\partial/\partial x)^m (\partial/\partial y)^n f(x, y)$	$(-iu)^m (-iv)^n F(u, v)$
(7) $x^m y^n f(x, y)$	$(-i)^{m+n} (\partial/\partial u)^m (\partial/\partial v)^n F(u, v)$
(8) $2\pi \delta(x, y)$	1
(9) $2\pi \delta(x-a, y-b)$	$e^{i(au+bv)}$
(10) $\pi \delta(x-a, y-b) + \pi \delta(x+a, y+b)$	$\cos(au+bv)$
(11) $\pi \delta(x-a, y-b) - \pi \delta(x+a, y+b)$	$i \sin(au+bv)$
(12) $e^{-\frac{1}{2}(x^2/a^2 + y^2/b^2)}$	$abe^{-\frac{1}{2}(a^2u^2 + b^2v^2)}$
(13) $(x/a)^m (y/b)^n e^{-\frac{1}{2}(x^2/a^2 + y^2/b^2)}$	$i^{m+n} H_m(au) H_n(bv) abe^{-\frac{1}{2}(a^2u^2 + b^2v^2)}$
(14) $(\alpha/a)^m (\beta/b)^n H_m(\alpha x/\alpha a) H_n(b_1 y/\beta b)$ $\times e^{-\frac{1}{2}(x^2/a^2 + y^2/b^2)}$	$i^{m+n} H_m(a_1 u) H_n(b_1 v) abe^{-\frac{1}{2}(a^2u^2 + b^2v^2)}$ where $a_1^2 = a^2 + \alpha^2, b_1^2 = b^2 + \beta^2$.
(15) $e^{-\frac{1}{2}[(\alpha x + \beta y)^2/a^2 + (\gamma x + \delta y)^2/b^2]}$ where $\alpha\delta - \beta\gamma = 1$	$abe^{-\frac{1}{2}[a^2(\delta u - \gamma v)^2 + b^2(-\beta u + \alpha v)^2]}$
(16) $(1 + x^2/a^2 + y^2/b^2)^{-3/2}$	$abe^{-\frac{1}{2}(a^2u^2 + b^2v^2)} \frac{1}{2}$
(17) 1, $ x < a, y < b$ 0, elsewhere	$4ab \frac{\sin au}{au} \frac{\sin bv}{bv}$
(18) $(1 - x /2a)(1 - y /2b),$ $ x < 2a$ $ y < 2b$ 0, elsewhere	$4ab \left[\frac{\sin au}{au} \frac{\sin bv}{bv} \right]^2$

then

$$f(x, y) = \Sigma \Sigma c_{mn} \varphi(a_m^2 - a^2, b_n^2 - b^2).$$

Again, if

$$g(x, y) = \varphi(a_1^2, b_1^2) \Sigma \Sigma c_{mn} (A_1 x)^m (B_1 y)^n, \quad \dots\dots\dots (8.3)$$

then

$$f(x, y) = \varphi(a_2^2, b_2^2) \Sigma \Sigma c_{mn} H_m(A_2 x) H_n(B_2 y),$$

where $a_2^2 = a_1^2 - a^2$, $b_2^2 = b_1^2 - b^2$; $A_1 = a_2/a_1a$, $B_1 = b_2/b_1b$, $A_2 = a_1/a_2a$, $B_2 = b_1/b_2b$. The last result can also be verified by direct integration of equation (2.1); variants of it have been discussed by Malmquist (1943), van de Hulst (1946), and Lyttkens (1949). Finally, if

$$g(x, y) = (ab)^{-1} e^{-\frac{1}{2}(x'^2/a_1^2 + y'^2/b_1^2)}, \quad \dots \dots \dots (8.4)$$

$$\begin{aligned} x' &= x \cos \theta + y \sin \theta, \\ y' &= -x \sin \theta + y \cos \theta, \end{aligned}$$

then

$$\begin{aligned} f(x, y) &= (AB)^{-1} e^{-\frac{1}{2}(X'^2/A^2 + Y'^2/B^2)}, \\ X' &= x \cos \Theta + y \sin \Theta, \\ Y' &= -x \sin \Theta + y \cos \Theta, \end{aligned}$$

where A, B, Θ are given by

$$\begin{aligned} \cot 2\Theta &= \cot 2\theta - \operatorname{cosec} 2\theta (a^2 - b^2)/(a_1^2 - b_1^2), \\ 2A^2 &= a_1^2 + b_1^2 - a^2 - b^2 + (a_1^2 - b_1^2) \sin 2\theta \operatorname{cosec} 2\Theta, \\ 2B^2 &= a_1^2 + b_1^2 - a^2 - b^2 + (b_1^2 - a_1^2) \sin 2\theta \operatorname{cosec} 2\Theta. \end{aligned}$$

We take the value of Θ which lies in the same quadrant as θ , and this value is always real and uniquely determined. But if the expression for A^2 or B^2 is negative, there is no solution.

When $g(x, y)$ is not expressible with sufficient accuracy in one of the forms (8.2), (8.3), (8.4), it may still be possible to express it as a sum of one of these functions and a residual, and the two components can be sharpened separately. If the residual and its gradient are numerically small, the sharpening of the residual by approximate methods is a much simpler task than the sharpening of the distribution $g(x, y)$.

Reverting to an arbitrary kernel, if $g(x, y)$ is a doubly periodic function whose Fourier series is

$$g(x, y) = \sum_m \sum_n [g_{mn} \cos (max + nby) + g'_{mn} \sin (max + nby)],$$

it is found by using Table 2 that

$$f(x, y) = \sum_m \sum_n [f_{mn} \cos (max + nby) + f'_{mn} \sin (max + nby)],$$

where

$$f_{mn} + if'_{mn} = (g_{mn} + ig'_{mn})/2\pi H(ma, nb).$$

When f, g, h all have circular symmetry, their Fourier transforms become Fourier-Bessel transforms, by the use of which several particular solutions of equation (2.1) may be obtained. Some Fourier-Bessel transforms which will be found useful for this purpose are listed in Table 3. The two columns of this table may be interchanged by interchanging r and s . Some of these transforms were taken from Watson (1944) and others were derived from them by using the rules on lines (1)–(6) of the table. The polynomials $Q_n(x)$ may be defined by

$$Q_n(x)e^{-x} = [(d/dx)(xd/dx)]^n e^{-x},$$

TABLE 3
FOURIER-BESSEL TRANSFORMS

J_0, J_1, K_0, K_1 denote Bessel functions in the notation of Watson (1944). ∇^2 is the Laplacian operator $s^{-1}(d/ds)(sd/ds)$, and Q_n, R_n are the polynomials

$$Q_n(x) = n! \sum_{p=0}^n (-1)^{n-p} x^p / p! (n-p)!,$$

$$R_n(x) = 2^{-n} \sum_{p=0}^n (2n-p)! (2x)^p / p! (n-p)!$$

$f(r) \equiv \int_0^\infty F(s) J_0(rs) s ds$	$F(s) \equiv \int_0^\infty f(r) J_0(rs) r dr$
(1) $af_1(r) + bf_2(r)$	$aF_1(s) + bF_2(s)$
(2) $f(r/a)$	$a^2 F(as)$
(3) $2\pi f_1(r) f_2(r)$	$\int_0^\infty \int_{-\pi}^\pi F_1(s') F_2(s) s' ds' d\theta,$ where $S^2 = s^2 + s'^2 - 2ss' \cos \theta$
(4) $\int_r^\infty xf(x) dx$	$-s^{-1} F'(s)$
(5) $(d/dr)\{rf(r)\}$	$-(d/ds)\{sF(s)\}$
(6) $r^{2n} f(r)$	$(-1)^n \nabla^{2n} F(s)$
(7) $2\pi \delta(r)$	1
(8) $2\pi \delta(r-a)$	$J_0(as)$
(9) $e^{-\frac{1}{2}r^2/a^2}$	$a^2 e^{-\frac{1}{2}a^2 s^2}$
(10) $(\frac{1}{2}r^2/a^2)^n e^{-\frac{1}{2}r^2/a^2}$	$(-1)^n Q_n(\frac{1}{2}a^2 s^2) a^2 e^{-\frac{1}{2}a^2 s^2}$
(11) $(\alpha/a)^{2n} Q_n(\frac{1}{2}\alpha^2 r^2/\alpha^2 a^2) e^{-\frac{1}{2}r^2/a^2}$	$(-1)^n Q_n(\frac{1}{2}\alpha^2 s^2) a^2 e^{-\frac{1}{2}a^2 s^2},$ where $\alpha_1^2 = a^2 + \alpha^2$
(12) $(a^2 + r^2)^{-\frac{1}{2}}$	$s^{-1} e^{-as}$
(13) $(a^2 + r^2)^{-3/2}$	$a^{-1} e^{-as}$
(14) $1.3.5 \dots (2n+1)(a^2 + r^2)^{-\frac{1}{2}(2n+3)}$	$R_n(as) a^{-2n-1} e^{-as}$
(15) $(a^2 + r^2)^{-1}$	$K_0(as)$
(16) $2a^2(a^2 + r^2)^{-2}$	$as K_1(as)$
(17) $4a^4(a^2 + r^2)^{-3}$	$as K_1(as) + \frac{1}{2} a^2 s^2 K_0(as)$
(18) $(\pi a^2)^{-1}, \quad r < a$ $0, \quad r > a$	$\frac{1}{2\pi} \left[\frac{2}{as} J_1(as) \right]$
(19) $2\pi^{-2} a^{-2} \{\arccos(r/2a) - (r/2a)\sqrt{1-(r/2a)^2}\}, \quad r < 2a$ $0, \quad r > 2a$	$\frac{1}{2\pi} \left[\frac{2}{as} J_1(as) \right]^2$

and the first five of them are

$$\begin{aligned} Q_0(x) &= 1, & Q_1(x) &= x-1, & Q_2(x) &= x^2-4x+2, \\ Q_3(x) &= x^3-9x^2+18x-6, & Q_4(x) &= x^4-16x^3+72x^2-96x+24. \end{aligned}$$

The first five of the polynomials $R_n(x)$ are

$$\begin{aligned} R_0(x) &= 1, & R_1(x) &= x+1, & R_2(x) &= x^2+3x+3, \\ R_3(x) &= x^3+6x^2+15x+15, & R_4(x) &= x^4+10x^3+45x^2+105x+105. \end{aligned}$$

Also

$$R_n(x) = (2n-1)R_{n-1}(x) + x^2 R_{n-2}(x).$$

For brevity, we write

$$\begin{aligned}\varphi(a^2) &= (2\pi a^2)^{-1} e^{-\frac{1}{2}r^2/a^2}, \\ \psi(a) &= (a/2\pi)(a^2 + r^2)^{-3/2}.\end{aligned}$$

Then, if $h(r) = \varphi(a^2)$ and $g(r) = \sum c_n \varphi(a_n^2)$, we find

$$f(r) = \sum c_n \varphi(a_n^2 - a^2).$$

This form of solution was known to Kapteyn (1920), who used it to compute true frequency distributions of stellar proper motions. With the same kernel, if

$$g(r) = \varphi(a_1^2) \sum c_n (\frac{1}{2} A_1^2 r^2)^n,$$

then

$$f(r) = \varphi(a_2^2) \sum c_n Q_n(\frac{1}{2} A_2^2 r^2)^n,$$

where $a_2^2 = a_1^2 - a^2$, $A_1 = a_2/a_1 a$, $A_2 = a_1/a_2 a$.

If $h(r) = \psi(a)$ and $g(r) = \sum c_n \psi(a_n)$, we find

$$f(r) = \sum c_n \psi(a_n - a).$$

The convolution of $\varphi(a^2)$ and $\psi(b)$ is the analogue of the Voigt function in one dimension.

Other particular solutions may be found from Table 3 when the kernel is of the form $(2\pi a^2)^{-1} e^{-r/a}$ or of the form $(4\pi a^3)^{-1} r K_1(r/a)$.

Fourier-Bessel transforms may also be computed by numerical integration. For best results the major variations of the function should be taken out by using the functions of Table 3, and the residual function transformed numerically. This process is tedious, but has been used successfully to sharpen the observed light curves of external galaxies with circular symmetry. Numerical integration of Fourier transforms has been used similarly in one dimension by Stokes (1948) and by Fellgett and Schmeidler (1952).

Instability of the kind described in Section IV arises through the uncertainty of $F(u, v)$ for large (u, v) . In practice, plausible solutions can be obtained by extrapolating $F(u, v)$ outside the region in which it is well determined, on the assumption that it falls "smoothly" to zero.

IX. APPROXIMATE INVERSE KERNELS

In this section, following a method introduced by Kreisel (1949), we overcome the instability of the solution by seeking only partially sharpened solutions $f_n(x, y)$, given by the convolution of the true solution $f(x, y)$ with pre-assigned smoothing functions $j_n(x, y)$ which are, loosely speaking, broad enough to smooth out spurious undulations of short wavelength in $f(x, y)$, but not so broad as to distort seriously the significant features of $f(x, y)$.

For brevity, we now write $2\pi H(u, v) = \bar{H}(u, v)$, and similarly for other functions of (u, v) . The arguments (x, y) and (u, v) following functional symbols

will be omitted when there is no risk of confusion, and the convolution of any two functions f_1, f_2 will be denoted by $f_1 * f_2$. Equation (2.1) is thus written

$$g = h * f. \quad \dots\dots\dots (9.1)$$

As always, we assume that h is L. If f is L also, then

$$\bar{G} = \overline{H\bar{F}}.$$

We define $\bar{K}(u, v)$ by $\bar{K}\bar{H} = 1$, so that

$$\bar{F} = \bar{K}\bar{G}.$$

(In this section, the symbols J, K do not refer to Bessel functions.) If $K(u, v)$ were the transform of a function $k(x, y)$ of L, the solution of (9.1) would therefore be $f = k * g$, and $k(x, y)$ could be called the inverse kernel. But if $h(x, y)$ is a proper function, we know that $K(u, v) \rightarrow \infty$ with $u^2 + v^2$, so that $k(x, y)$ does not exist.

However, in many cases there exists a sequence of approximate inverse kernels $k_n(x, y)$ of L, and associated smoothing functions $j_n = k_n * h$, such that the approximate solution $f_n = k_n * g$ is the smooth of f with j_n , is stable, and under certain conditions the sequence $f_n(x, y)$ converges uniformly to $f(x, y)$. Observe that, when $g(x, y)$ is specified by its values at a finite set of points together with an interpolation formula, there is in general no exact solution $f(x, y)$, and the best we can do is to find an approximate solution f_n such that $g_n = h * f_n$ agrees with the observational data within acceptable limits of error. It will be shown that, whether $f(x, y)$ exists or not, the sequence $g_n(x, y)$ converges uniformly to $g(x, y)$; provided that $g(x, y)$ is bounded and of bounded gradient everywhere, that is,

$$\begin{aligned} |g(x', y') - g(x, y)| &\leq m\varphi(x' - x, y' - y), \quad \dots\dots\dots (9.2) \\ \varphi(x, y) &= |(x^2 + y^2)^{\frac{1}{2}}|, \quad x^2 + y^2 \leq M^2/m^2, \\ &= M/m, \quad x^2 + y^2 > M^2/m^2; \end{aligned}$$

and provided also that, for $\varepsilon > 0$, there exists an integer N such that

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |j_n(x, y)| \varphi(x, y) dx dy < \varepsilon/m \quad \dots\dots\dots (9.3)$$

for every $n \geq N$. We suppose also that the $j_n(x, y)$ are normalized :

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} j_n(x, y) dx dy = 1, \quad n = 1, 2, 3, \dots\dots\dots (9.4)$$

We now give a formal statement and proof of these results.

Statement.—If

- (i) g is bounded, and integrable in every finite region,
- (ii) g is of bounded gradient everywhere, as in (9.2),
- (iii) functions k_n and $j_n = k_n * h$ can be chosen such that (9.3) and (9.4) hold, and h, k_n are L,
- (iv) $f_n = k_n * g$;

then

- (v) f_n is bounded, and integrable in every finite region, so that $g_n = h * f_n$ exists,
- (vi) j_n is L, and $g_n = j_n * g$,
- (vii) f_n is continuously determined by g ,
- (viii) $|g_n - g| \rightarrow 0$ everywhere uniformly as $n \rightarrow \infty$.

If also

- (ix) f exists such that $g = h * f$, and f is bounded, and integrable in every finite region;

then

- (x) $f_n = j_n * f$.

If also

- (xi) f is of bounded gradient everywhere;

then

- (xii) $|f_n - f| \rightarrow 0$ everywhere uniformly as $n \rightarrow \infty$.

Proofs

- (v) Since $f_n = k_n * g$, and k_n is L, (v) follows from (i).

(vi) Since h, k_n are L and $j_n = k_n * h$, therefore j_n is L. Also, by (v), (iv), and (iii), $g_n = h * (k_n * g) = (h * k_n) * g = j_n * g$, the change in order of integration being justified by absolute convergence.

(vii) Let $\Delta g(x, y)$ be a variation of $g(x, y)$, such that $|\Delta g(x, y)| < \varepsilon$ for all (x, y) . Let $\Delta f_n(x, y)$ be the corresponding variation of $f_n(x, y)$. Then, since k_n is L,

$$|\Delta f_n| = |k_n * \Delta g| \leq |k_n| * |\Delta g| < \varepsilon \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |k_n(x, y)| \, dx dy = \varepsilon C_n,$$

where C_n is a positive number depending only on n . Therefore $f_n(x, y)$ is continuously determined by $g(x, y)$, that is, it is stable. (But note that this continuity is not uniform with respect to n , for $C_n \rightarrow \infty$ with n .)

- (viii) From (vi) and (9.4),

$$\begin{aligned} |g_n(x', y') - g(x', y')| &= \left| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} j_n(x' - x, y' - y) [g(x, y) - g(x', y')] \, dx dy \right| \\ &\leq m \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |j_n(x' - x, y' - y)| \, \varphi(x' - x, y' - y) \, dx dy \\ &< \varepsilon, \text{ for } n \geq N, \end{aligned}$$

by (9.2) and (9.3).

- (x) Since h, k_n are L, we have from (iv) and (ix),

$$f_n = k_n * (h * f) = (k_n * h) * f = j_n * f.$$

- (xii) Proof is similar to the proof of (viii).

The treatment above owes much to Kreisel's "Theorem A" (1949). In our notation, Kreisel assumes (ii), (iii), (ix), and (x) and deduces (iv), (vii), and in effect (viii), though without explicitly using the idea of a sequence. To facilitate comparison of the two treatments, we give a conversion table for notation:

Burr:	f	f_n	g	g_n	h	H	j_n	J_n	k_n	K_n
Kreisel:	ψ	$\bar{\psi}$	k	\bar{k}	L	Λ	μ	M	λ	M/Λ

The conditions (iii) may be relaxed sufficiently to admit improper functions of the form

$$k_n(x, y) = C_n \delta(x, y) + l_n(x, y), \quad \dots \quad (9.5)$$

where l_n belongs to L; and the theorem remains true, with slight modifications in the proofs.

In another variant, (9.3) is replaced by the condition that

$$\bar{J}_n(u, v) \rightarrow 1 \text{ uniformly in every finite region, } \dots \quad (9.6)$$

and (ii) is replaced by the condition that g is L and continuous, with the remaining conditions unchanged. It follows that g_n is L and continuous, and $\bar{G}_n = \bar{J}_n \bar{G}$, and proposition (viii) is replaced by

$$\begin{aligned} \bar{G} - \bar{G}_n &= (1 - \bar{J}_n) \bar{G} \\ &\rightarrow 0 \text{ as } n \rightarrow \infty, \quad \dots \quad (9.6a) \end{aligned}$$

by (9.6). In this variant, the writer has been unable to prove that $g - g_n \rightarrow 0$ except in the rather restrictive case in which g, \bar{G} both belong to L. However, most computers will accept (9.6a) as sufficient justification for using the process with any observational data. Similar remarks apply to f, \bar{F} when these functions exist. The proofs of (v), (vi), (vii), and (x) are unchanged.

We now give two forms for the sequence of smoothing functions which have been used in numerical problems.

Kreisel (1949) chooses smoothing functions of the form

$$j_n(x, y) = (2\pi a_n b_n)^{-1} e^{-\frac{1}{2}(x^2/a_n^2 + y^2/b_n^2)}, \quad \dots \quad (9.7)$$

where $a_n, b_n \rightarrow 0$, as $n \rightarrow \infty$, so that (9.3) and (9.4) are satisfied. The approximate inverse kernels $k_n(x, y)$ have then to be computed through their Fourier transforms:

$$\bar{K}_n(u, v) = [\bar{H}(u, v)]^{-1} e^{-\frac{1}{2}(a_n^2 u^2 + b_n^2 v^2)}. \quad \dots \quad (9.8)$$

The method is applicable only if the $k_n(x, y)$ exist and belong to L. It is valid for algebraic kernels such as (2.7), for which some functions $k_n(r)$ have been tabulated by Bullard and Cooper (1948). But, if $H(u, v)$ has zeros or if the kernel is Gaussian, the form (9.7) is not applicable.

The second form for the sequence is given by

$$\bar{J}_n = 1 - (1 - \bar{H})^n, \quad \dots \quad (9.9.)$$

$$\begin{aligned} \bar{K}_n &= 1 + (1 - \bar{H}) + (1 - \bar{H})^2 + \dots + (1 - \bar{H})^{n-1} \\ &= {}_n c_2 \bar{H} + {}_n c_3 \bar{H}^2 - \dots + (-\bar{H})^{n-1}. \quad \dots \quad (9.10) \end{aligned}$$

Since $\bar{H}^2, \bar{H}^3, \dots$ are the Fourier transforms of $2\pi h * h, 2\pi h * h * h, \dots$, which belong to L , it follows that j_n is L , and that

$$k_n(x, y) = n\delta(x, y) - l_n(x, y), \dots\dots\dots (9.11)$$

where $l_n(x, y)$ is L . Also, from (9.9), since $\bar{H}(u, v)$ is continuous, we see that (9.6) is satisfied provided that

$$|1 - \bar{H}(u, v)| < 1 \dots\dots\dots (9.12)$$

everywhere. We conclude that when the kernel satisfies (9.12) the method is applicable under the conditions discussed under equation (9.6a). Approximate solutions of this form have been discussed or applied by several authors, for example Burger and van Cittert (1932, 1933) and van de Hulst (1941).

It is easily verified from (9.9) and (9.10) that the sequence f_n satisfies the recurrence relations

$$f_{n+1} = g + f_n - h * f_n, \quad n = 0, 1, 2, \dots, \dots (9.13)$$

$$f_{n+1} - f_n = (f_n - f_{n-1}) - h * (f_n - f_{n-1}), \quad n = 1, 2, 3, \dots, \dots (9.14)$$

with $f_0 = 0, f_1 = g$. From (9.13) we can compute f_2, f_3, f_4, \dots in turn, one integration being required at each step. (By one integration we mean one for each point (x, y) .) Note that an error $e(x, y)$ introduced at any step will be reduced after a further m integrations to an error whose Fourier transform is $[1 - \bar{H}(u, v)]^m E(u, v)$, which tends to zero as $m \rightarrow \infty$. Equation (9.13) may also be written $f_{n+1} - f_n = g - g_n$, so that the process may be stopped when $f_{n+1} - f_n$ is smaller than the errors of observation in g . The numerical integration is easier if (9.14) is used instead of (9.13), because $f_n - f_{n-1}$ is in general numerically smaller and of smaller gradient than f_n . We have taken for our first approximation $f_1 = g$, but any function could have been taken. Thus the computer can use his judgement in selecting a trial solution for f_1 , compute f_2 from (9.13) and f_3, f_4, \dots from (9.13) or (9.14). In numerical examples with circular symmetry, it was found that graphs of the sequence $(f_{n+1} - f_n)$ along a diameter formed a "smooth" sequence of curves, and after plotting four or five such curves it was possible to extrapolate the sequence with tolerable accuracy and sketch the form of the next two or three curves. The sum of the ordinates of all these curves gave the distribution finally adopted for $f(x, y)$.

Alternatively we may compute the approximate inverse kernel k_n in the form (9.11) for some particular n , say $n = 5$, and then obtain f_n in a single integration:

$$f_n = k_n * g \equiv ng - l_n * g. \dots\dots\dots (9.15)$$

We then form $g_n = h * f_n$, and, if $g - g_n$ is not negligible, we can compute f_{2n}, f_{3n}, \dots in turn using the recurrence relation

$$f_{(m+1)n} - f_{mn} = k_n * (g - g_{mn}), \quad m = 0, 1, 2, \dots, \dots (9.16)$$

in which we have taken $f_0 = 0$, although it is possible to choose f_0 arbitrarily, setting $g_0 = h * f_0$. The computation of $l_n(x, y)$ is tedious, but is worth while if

several observed distributions with the same kernel have to be sharpened. It can be computed by $n-2$ integrations,

$$l_n = {}_n c_2 h - {}_n c_3 h * h + {}_n c_4 h * h * h - \dots,$$

or by taking two Fourier transforms :

$$\bar{L}_n = n - \frac{1 - (1 - \bar{H})^n}{\bar{H}},$$

though this may not be practicable except in cases with circular symmetry. If the kernel is expressible as a sum of Gaussian terms or as a sum of terms of the form $a(b^2 + r^2)^{-3/2}$, then $l_n(r)$ can be expressed in the same form. For example, if

$$h(x, y) = \varphi(a^2, b^2) \equiv (2\pi ab)^{-1} e^{-\frac{1}{2}(x^2/a^2 + y^2/b^2)},$$

then

$$l_5(x, y) = 10\varphi(a^2, b^2) - 10\varphi(2a^2, 2b^2) + 5\varphi(3a^2, 3b^2) - \varphi(4a^2, 4b^2).$$

This method has been used to correct light curves of the same external galaxies as those mentioned in the preceding section. The function $l_{10}(r)$ corresponding to the kernel (2.3) was computed, and it was found possible to integrate $l_{10} * g$ with sufficient accuracy to determine $f_{10}(r)$ within about 2 per cent. The agreement of the results obtained by the two methods was very satisfying.

From (9.15), it appears that random errors in the data $g(x, y)$ will be magnified at least n times in $f_n(x, y)$. Thus the solution becomes increasingly unstable as n is increased. Significant errors may also be introduced by repeated numerical integrations. But by using 50 or more cells on the transparent grid representing $h(x, y)$ or $l_n(x, y)$, these errors may be kept reasonably small. This has been verified by a test in which $h(r) = (2\pi)^{-1} e^{-\frac{1}{2}r^2}$, $f(r) = 200e^{-\frac{1}{2}r^2}$, $g(r) = 100e^{-r^2/4}$. The approximations f_2, f_3, f_4, f_5 were computed analytically, and also by four numerical integrations using (9.14) with a grid of 55 cells representing $h(r)$. The errors due to numerical integration in f_2, f_3, f_4, f_5 were found to be not greater than 0.6, 1.0, 1.2, and 1.5 respectively. $f_5(r)$ was also computed directly from (9.15) with a grid of 79 cells representing $l_5(r)$. The errors in this case were found to be less than 1.0. These errors are regarded as reasonably small, because they are not larger than the errors of observation to be expected in $g(r)$, which are usually at least 1 or 2 per cent. of $g(0)$.

X. CONCLUSIONS

It is not possible to lay down fixed rules for deciding which method of solution is best for a particular problem, but the following may serve as a rough guide.

(a) If the corrections are small and the kernel has the requisite moments, use a polynomial solution.

(b) If the corrections are large, or the moments do not exist, or both, whilst $|H(u, v)| > 0$ everywhere, then Kreisel's method or the solutions (9.13)–(9.16) (or both) may be applicable.

(c) If $H(u, v) = 0$ in some region but $|1 - \bar{H}(u, v)| < 1$ elsewhere, relations (9.13)–(9.16) will give approximations to the principal solution, provided that the unwanted Fourier components are filtered out as described by Bracewell and Roberts (1954).

(d) If (a), (b), and (c) are inapplicable, the solution must be obtained by Fourier transforms or by trial and error. In problems with circular symmetry, the use of Fourier-Bessel transforms is sometimes preferable to the methods in (b) and (c).

XI. ACKNOWLEDGMENTS

The author is indebted to Dr. S. Gascoigne and Dr. G. de Vaucouleurs for encouragement and helpful criticism.

XII. REFERENCES

- BOCHNER, S. (1932).—"Vorlesungen über Fouriersche Integrale." (Akademische Verlagsgesellschaft: Leipzig.)
- BODE, H. W., and SHANNON, C. E. (1950).—*Proc. Inst. Radio Engrs.*, N.Y. **38**: 417.
- BOLTON, J. G., and WESTFOLD, K. C. (1950).—*Aust. J. Sci. Res. A* **3**: 25.
- BRACEWELL, R. N., and ROBERTS, J. A. (1954).—*Aust. J. Phys.* **7**: 615.
- BULLARD, E. C., and COOPER, R. I. B. (1948).—*Proc. Roy. Soc. A* **194**: 332.
- BURGER, H. C., and VAN CITTERT, P. H. (1932).—*Z. Phys.* **79**: 722.
- BURGER, H. C., and VAN CITTERT, P. H. (1933).—*Z. Phys.* **81**: 428.
- COUTREZ, R. (1949).—*Commun. Obs. Belg.* No. 11.
- EDDINGTON, A. S. (1940).—*Mon. Not. R. Astr. Soc.* **100**: 354.
- FELLGETT, P. B., and SCHMEIDLER, F. B. (1952).—*Mon. Not. R. Astr. Soc.* **112**: 445.
- VAN DE HULST, H. C. (1941).—*B.A.N.* **9**: 225.
- VAN DE HULST, H. C. (1946).—*B.A.N.* **10**: 75.
- KAPTEYN, J. C. (1920).—*Publ. Astr. Lab. (Obs.) Groningen* **30**: 43.
- KREISEL, G. (1949).—*Proc. Roy. Soc. A* **197**: 160.
- LYTTKENS, E. (1949).—*Ark. Astr.* **1**, Nos. 3, 4, 5, 7.
- MALMQUIST, K. G. (1943).—*Ark. Mat. Astr. Fys.* **29B**, No. 8.
- STOKES, A. R. (1948).—*Proc. Phys. Soc. Lond.* **61**: 382.
- DE VAUCOULEURS, G. (1948).—*Ann. Astrophys.* **11**: 247.
- WATSON, G. N. (1944).—"Theory of Bessel Functions." 2nd Ed. (Cambridge Univ. Press.)

CORRECTING FOR GAUSSIAN AERIAL SMOOTHING

By R. N. BRACEWELL*

[*Manuscript received August 25, 1954*]

Summary

Let a two-dimensional survey with a Gaussian aerial beam establish values at intervals of $\sqrt{2}$ standard deviations. Then the correction for aerial smoothing is simply calculated as the difference between the value to be corrected and the mean of the neighbouring four values.

I. INTRODUCTION

An important part of the reduction of radio-astronomical observations consists in correcting for the blurring which arises from the finite extent of the aerial beam. Methods for doing this in the one-dimensional case have been discussed by Bracewell and Roberts (1954), and in principle, and indeed in practice, the methods may be extended to two dimensions. It was shown, however, that some of the detail is irretrievably lost; and, as there is considerable labour involved in calculating the correction, which is in any case only partial, there has been a tendency in recent publications of two-dimensional observations to omit any correction for aerial smoothing (e.g. McGee and Bolton 1954).

There appears therefore to be scope for new methods of correction involving procedures less elaborate than the calculation of two-dimensional convolutions or two-dimensional Fourier transforms. A sacrifice of accuracy would seem reasonable in order to obtain this simplicity, especially where the correction to be applied is small. A simple, approximate method of correction would be valuable for quickly seeing the effect of the correction and for deciding whether the labour of calculating a small correction would be worth while.

In some current researches, the labour involved in applying the established methods is already prohibitive as a result of the vast increase in the mass of observational data yielded by the high-resolution aerals employed. The Mills radiometer (Mills and Little 1953) is a case in point. An attempt has therefore been made to find a simple, approximate two-dimensional method of correction for a Gaussian aerial beam as used in that case.

The approach to the problem was to list simple numerical operations on a two-dimensional array of data (graphical or tabular) and to explore the possibilities of each in turn. The operation of taking finite differences is the one which has proved fruitful. It is applicable to aerial beams other than Gaussian.

As a by-product, the present theory provides, in the one-dimensional case, a sound alternative to a formula given by Eddington (1913) for correcting the blurring effect of a Gaussian distribution of errors. Eddington's series, which

* Division of Radiophysics, C.S.I.R.O., University Grounds, Sydney.

is widely known and used in astronomy (though not apparently in radio astronomy), has proved valuable in practice, but suffers from being theoretically doubtful (Jeffreys 1938) and from the fact that in actual use the series of derivatives is evaluated as a series of finite differences.

II. THEORY

Some preliminary notes are needed on the theory of finite differences in relation to Fourier theory. If from a given function another is derived by taking finite differences, then its spectrum is related to that of the given function in a way which will now be shown; first for functions of one variable, and then for functions of two variables.

Let

$${}^{\alpha}\Delta f(x) \equiv f(x + \tfrac{1}{2}\alpha) - f(x - \tfrac{1}{2}\alpha),$$

i.e. ${}^{\alpha}\Delta f(x)$ represents the result of taking the difference between two values of $f(x)$, at values of x separated by an interval α .

Now the Fourier transform of $f(x+x_0)$ is, by the shift theorem, $\exp(2\pi i x_0 s) \bar{f}(s)$, where $\bar{f}(s)$ is the Fourier transform of $f(x)$. Hence

$$\begin{aligned} \text{F.T. of } {}^{\alpha}\Delta f(x) &= e^{\pi i \alpha s} \bar{f}(s) - e^{-\pi i \alpha s} \bar{f}(s) \\ &= 2i \sin(\pi \alpha s) \bar{f}(s). \end{aligned}$$

Thus, the effect of differencing a function of one variable at interval α is to multiply the spectrum by $2i \sin \pi \alpha s$.

In the two-dimensional case let $f(x, y)$ have a two-dimensional Fourier transform $\bar{f}(s_x, s_y)$. For differences at interval α in the x -direction we write

$${}^{\alpha}\Delta_x f(x, y) = f(x + \tfrac{1}{2}\alpha, y) - f(x - \tfrac{1}{2}\alpha, y),$$

and for differences at interval β in the y -direction

$${}^{\beta}\Delta_y f(x, y) = f(x, y + \tfrac{1}{2}\beta) - f(x, y - \tfrac{1}{2}\beta).$$

As an example of the notation for differences of higher order,

$$\begin{aligned} {}^{\alpha\alpha}\Delta_{xx} f(x, y) &= f(x + \tfrac{1}{2}\alpha, y) - 2f(x, y) + f(x - \tfrac{1}{2}\alpha, y), \\ {}^{\alpha\beta}\Delta_{xy} f(x, y) &= f(x + \tfrac{1}{2}\alpha, y + \tfrac{1}{2}\beta) - f(x - \tfrac{1}{2}\alpha, y + \tfrac{1}{2}\beta) \\ &\quad + f(x - \tfrac{1}{2}\alpha, y - \tfrac{1}{2}\beta) - f(x + \tfrac{1}{2}\alpha, y - \tfrac{1}{2}\beta). \end{aligned}$$

By the two-dimensional shift theorem, the two-dimensional Fourier transform of $f(x+x_0, y+y_0)$ is

$$\exp[i2\pi(x_0 s_x + y_0 s_y)] \bar{f}(s_x, s_y).$$

Hence

$$\begin{aligned} \text{2-dim. F.T. of } {}^{\alpha}\Delta_x f(x, y) &= e^{\pi i \alpha s_x} \bar{f}(s_x, s_y) - e^{-\pi i \alpha s_x} \bar{f}(s_x, s_y) \\ &= 2i \sin(\pi \alpha s_x) \bar{f}(s_x, s_y), \end{aligned}$$

$$\text{2-dim. F.T. of } {}^{\beta}\Delta_y f(x, y) = 2i \sin(\pi \beta s_y) \bar{f}(s_x, s_y),$$

$$,, ,, {}^{\alpha\alpha}\Delta_{xx} f(x, y) = (2i \sin \pi \alpha s_x)^2 \bar{f}(s_x, s_y),$$

$$,, ,, {}^{\alpha\beta}\Delta_{xy} f(x, y) = (2i \sin \pi \alpha s_x)(2i \sin \pi \beta s_y) \bar{f}(s_x, s_y).$$

With these necessary properties established, we turn to the study of the integral equation

$$T_a(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(x-u, y-v) T(u, v) du dv,$$

where $T(x, y)$ is the true distribution of temperature and $T_a(x, y)$ is the observed distribution. To represent a Gaussian aerial pattern we take

$$A(x, y) = \frac{1}{2\pi\sigma\tau} \exp\left\{-\left(\frac{x^2}{2\sigma^2} + \frac{y^2}{2\tau^2}\right)\right\},$$

where σ and τ are the standard deviations in the x and y directions respectively.

Given T_a and A it is now required to deduce T . By the two-dimensional convolution theorem

$$\bar{T}_a(s_x, s_y) = \bar{A}(s_x, s_y) \bar{T}(s_x, s_y),$$

where \bar{T}_a , \bar{A} , and \bar{T} are the two-dimensional Fourier transforms of T_a , A , and T respectively. Since

$$\bar{A}(s_x, s_y) = \exp\{-2\pi^2(\sigma^2 s_x^2 + \tau^2 s_y^2)\},$$

it follows that

$$\bar{T}(s_x, s_y) = \exp\{2\pi^2(\sigma^2 s_x^2 + \tau^2 s_y^2)\} \bar{T}_a(s_x, s_y).$$

Now using the formula

$$\exp \theta^2 = 1 + \sin^2 \theta + \frac{5}{6} \sin^4 \theta + \frac{61}{90} \sin^6 \theta + \dots, \quad \theta^2 < \left(\frac{\pi}{2}\right)^2,$$

we have

$$\begin{aligned} \bar{T}(s_x, s_y) &= \left(1 + \sin^2 \pi \alpha s_x + \frac{5}{6} \sin^4 \pi \alpha s_x + \dots\right) \\ &\quad \times \left(1 + \sin^2 \pi \beta s_y + \frac{5}{6} \sin^4 \pi \beta s_y + \dots\right) \bar{T}_a(s_x, s_y) \\ &= \left(1 + \sin^2 \pi \alpha s_x + \sin^2 \pi \beta s_y + \frac{5}{6} \sin^4 \pi \alpha s_x \right. \\ &\quad \left. + \frac{5}{6} \sin^4 \pi \beta s_y + \sin^2 \pi \alpha s_x \sin^2 \pi \beta s_y + \dots\right) \bar{T}_a(s_x, s_y), \end{aligned}$$

where $\alpha = \sqrt{2}\sigma$, $\beta = \sqrt{2}\tau$, $\alpha^2 s_x^2 < \frac{1}{4}$, and $\beta^2 s_y^2 < \frac{1}{4}$.

Hence, taking transforms, and provided that $\bar{T}(s_x, s_y)$ does not extend outside the central rectangle of breadth α^{-1} and height β^{-1} ,

$$\begin{aligned} T(x, y) &= T_a(x, y) - \frac{1}{4} \alpha \alpha \Delta_{xx} T_a(x, y) - \frac{1}{4} \beta \beta \Delta_{yy} T_a(x, y) \\ &\quad + \frac{5}{96} \alpha \alpha \alpha \alpha \Delta_{xxxx} T_a(x, y) + \frac{5}{96} \beta \beta \beta \beta \Delta_{yyyy} T_a(x, y) \\ &\quad + \frac{1}{16} \alpha \alpha \beta \beta \Delta_{xxyy} T_a(x, y) + \dots \end{aligned}$$

This equation states that it is possible to obtain the true distribution by the application to the observed distribution of a series of corrections formed by

differencing the observations. If $T(x, y)$ contains Fourier components of spatial frequencies for which $|s_x| > \frac{1}{2}\alpha^{-1}$ or $|s_y| > \frac{1}{2}\beta^{-1}$ they will be severely reduced by Gaussian smoothing and will not be recoverable by the present process; however, their presence will not interfere with the application of the method. An approximately Gaussian aerial beam produced with a finite array will not receive spatial frequencies beyond a certain limit, and in the case of Mills's aerial this limit agrees, within a few per cent., with the natural limit set by the method.

Bearing in mind the purpose for which the equation was established, we now limit attention to the formulae resulting from the retention of two terms only. Then

$$\begin{aligned} T(x, y) &\simeq T_a(x, y) - \frac{1}{4} [T_a(x + \frac{1}{2}\alpha, y) + T_a(x - \frac{1}{2}\alpha, y) - 2T_a(x, y)] \\ &\quad - \frac{1}{4} [T_a(x, y + \frac{1}{2}\beta) + T_a(x, y - \frac{1}{2}\beta) - 2T_a(x, y)] \\ &= T_a(x, y) + \left[T_a(x, y) - \frac{T_a(x + \frac{1}{2}\alpha, y) + T_a(x - \frac{1}{2}\alpha, y) + T_a(x, y + \frac{1}{2}\beta) + T_a(x, y - \frac{1}{2}\beta)}{4} \right]. \end{aligned} \quad \dots\dots\dots (1)$$

$$= T_a(x, y) + \frac{[T_a(x, y) - T_a(x + \frac{1}{2}\alpha, y)] + [T_a(x, y) - T_a(x - \frac{1}{2}\alpha, y)]}{4} \\ + \frac{[T_a(x, y) - T_a(x, y + \frac{1}{2}\beta)] + [T_a(x, y) - T_a(x, y - \frac{1}{2}\beta)]}{4}. \quad \dots\dots\dots (2)$$

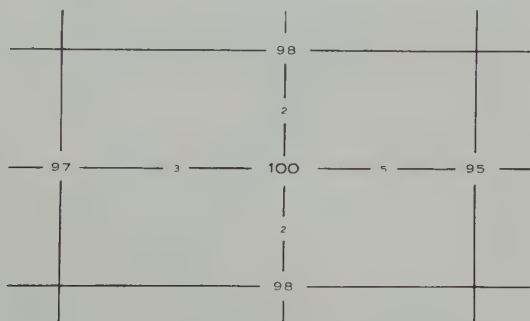


Fig. 1.—How the correction to the value 100 is calculated.

The correcting procedure based on formula (1) is as follows. *Tabulate the values of $T_a(x, y)$ in a two-dimensional array at intervals of $\sqrt{2}\sigma$ and $\sqrt{2}\tau$ in the x and y directions respectively. Subtract the mean of the four values surrounding any particular point from the value at that point, and the result is the correction to be applied to that value.* For example, in Figure 1 the mean of the four values surrounding the value 100 is 97. The corrected value is therefore 103. The alternative arrangement of formula (2) gives the correction as the mean excess of the value 100 over the four surrounding values, namely, $\frac{1}{4}(5+2+3+2)=3$.

III. TRIAL OF THE FORMULA

It is sufficient to test the formula in the one-dimensional case, where it reduces to

$$T(x) = T_a(x) - \frac{1}{4} \alpha \alpha \Delta_{xx} T_a(x).$$

A convenient distribution T in Figure 2 was taken and smoothed with a Gaussian aerial pattern to obtain the T_a shown. For comparison purposes, T_a was then smoothed so that the dotted curve, representing a single stage of restoration by

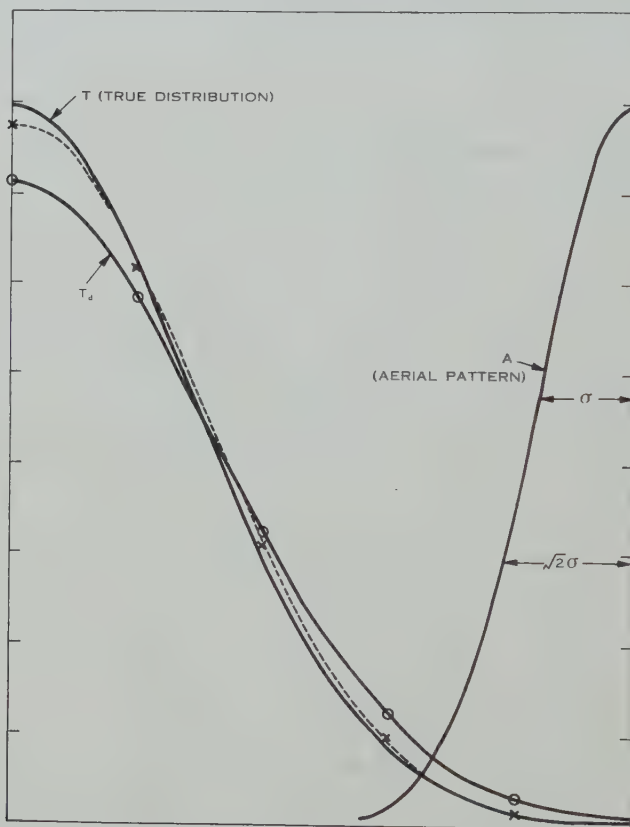


Fig. 2.—Numerical trial of correction by one stage of differencing.

- Values of T_a used for differencing.
- × Values corrected by differencing.
- One stage of correction by successive substitutions.

successive substitutions, could be plotted. T_a was then read off at the six points marked by circles, and the above formula applied to obtain the corrected values shown as crosses.

This test shows that, in this case, the simple formula gives the order of magnitude of the correction correctly and that it is as good as a single stage of correction by successive substitutions, which is of course much harder to compute, even in the one-dimensional case.

IV. RELATION TO EDDINGTON'S FORMULA

The one-dimensional solution is

$$T(x) = T_a(x) - \frac{1}{4}\Delta^2 T_a(x) + \frac{5}{96}\Delta^4 T_a(x) + \dots$$

This equation may be compared with Eddington's formula for dealing with the same problem

$$T(x) = T_a(x) - \frac{1}{4}T_a''(x) + \frac{1}{32}T_a'''(x) - \dots$$

In practical applications the differential coefficients are replaced by finite differences and, if only one correction term is taken, the two procedures may be shown to be the same since the Gaussian function assumed for the purposes of Eddington's formula has a standard deviation of $1/\sqrt{2}$.

The derivation of Eddington's formula is known to be doubtful (see Jeffreys 1938). The method of deriving a finite-difference formula by Fourier methods seems to be a better approach.

V. DISCUSSION

In approximate terms we may say that the negative second difference measures the upward convexity (or downward curvature) of the data. Where this is a maximum, as in the vicinity of a peak, the effect of the correction is greatest and is in such a sense as to increase the height of the peak. Near points of inflection the correction is small and changes sign. These features may be observed on Figure 3 which shows two curves, one obtained from the other by smoothing. The correction clearly agrees closely in sense and magnitude with the curvature.

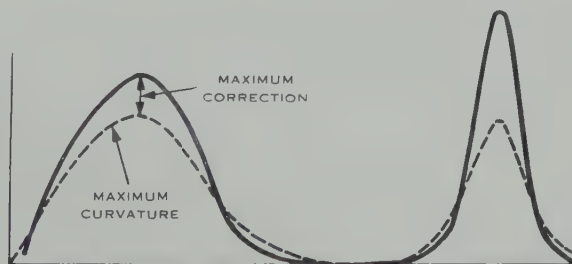


Fig. 3.—A curve (---) obtained from another (—) by smoothing.

The precise interval over which the differences are to be taken is prescribed by the theory. But if the data do not show much change in distances of several beam widths, then a coarser interval may be desirable to save numerical work. Also, small adjustments to the interval may be desired for numerical or graphical reasons. Differences taken over an interval m times wider than prescribed (ma instead of a) will be approximately m^2 times greater, since $\sin^2 \pi mas \simeq m^2 \sin^2 \pi as$. Within the limitations of this approximation, adjustments may be made to the interval, the factor m^2 being allowed for.

VI. REFERENCES

- BRACEWELL, R. N., and ROBERTS, J. A.—*Aust. J. Phys.* **7**: 615.
EDDINGTON, A. S. (1913).—*Mon. Not. R. Astr. Soc.* **73**: 359.
JEFFREYS, H. (1938).—*Mon. Not. R. Astr. Soc.* **98**: 190.
MCGEE, R. X., and BOLTON, J. G. (1954).—*Nature* **173**: 985.
MILLS, B. Y., and LITTLE, A. G. (1953).—*Aust. J. Sci. Res.* **6**: 272.

A METHOD OF CORRECTING THE BROADENING OF X-RAY LINE PROFILES

By R. N. BRACEWELL*

[*Manuscript received September 8, 1954*]

Summary

A numerical method is presented which allows X-ray line profiles to be corrected for instrumental broadening. It provides an alternative to the use of Fourier analysis as described by Stokes. A special development of the method (as hitherto used) is necessary to permit its application to X-ray analysis where the $K\alpha$ doublet constitutes a complication.

I. INTRODUCTION

X-ray powder photographs taken with characteristic X-radiation reveal facts about the powdered material, such as size of particle and condition of the crystal lattice structure, which have given rise to the developing subject of X-ray metallography (see Taylor 1945). The information is contained in the shape of the line profile as measured from the X-ray photograph with a microphotometer or directly with a Geiger counter spectrometer. However, the profile is subject to influences other than that under study and it is necessary to make allowance. For example, if the size of the particles is of interest, allowance must be made for the geometrical details of the X-ray camera and a number of other factors which combine to produce appreciable broadening; and if the state of strain of the lattice is under study then it is necessary to allow, in addition, for the broadening due to the finite size of the particle. Furthermore, $K\alpha$ radiation, commonly used in X-ray powder diffraction, is in fact a doublet which causes further complication when conditions are such that it is neither fully resolved nor fully unresolved.

In early work Scherrer corrected the width of the observed line by subtracting the width obtained when the phenomenon under study was absent. This is hardly reasonable (although there is a special case in which it is valid), and subsequently Taylor (1941) considered the modifications which arise if the line profiles are Gaussian. In this work the observations were not deemed adequate to yield more than one parameter, the line width. It has since become desirable to calculate the line profiles without initial assumptions as to shape. The Fourier analysis method for this has been described by Stokes (1948) and a relaxation method has been given by Paterson (1950).

The basic problem arises in very many connexions and has been recently discussed, with respect particularly to radio astronomy, by Bracewell and

* Division of Radiophysics, C.S.I.R.O., University Grounds, Sydney.

Roberts (1954). The process of successive substitutions, which they describe, ought to be applicable in X-ray analysis, bringing with it its advantages. Direct application is, however, usually impossible since the splitting of the $K\alpha$ line causes the process to diverge. This difficulty has been overcome by the introduction of a simple method of removing the effect of the doublet.

The present method is free from the subjective element of Paterson's method and rather shorter. It is much less laborious than Stokes's method. The latter method is, however, the appropriate one where, as in lattice distortion studies (cf. Averbach and Warren 1949), one requires the Fourier transform of the corrected profile.

II. THE PROBLEM

Let $A(\xi)$ be the apparatus function or instrumental profile. In the typical case shown in Figure 1, $A(\xi)$ has two humps corresponding to the two components of the doublet, and the broadening of each component is due to the geometry of the camera, etc. This profile was determined experimentally on annealed aluminium. If the aluminium is now subjected to cold work which would broaden an infinitely narrow line into a profile $F(\xi)$, then the distribution actually observed would be $G(\xi)$.

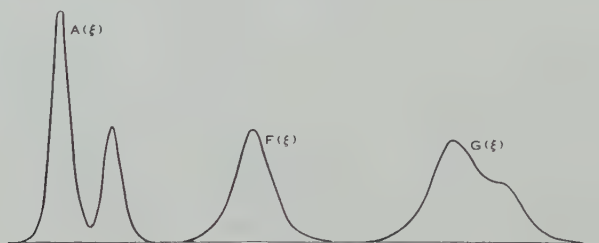


Fig. 1

The actual case illustrated in Figure 1 and tabulated in Table 1, has been taken from Paterson (1950), with small changes in normalization and interval of tabulation.

The relation between the three functions is

$$G(\xi) = \int_{-\infty}^{\infty} A(\xi - u)F(u)du. \quad \dots\dots\dots (1)$$

Given $G(\xi)$, and knowing the effect of the apparatus $A(\xi)$, it is required to find $F(\xi)$.

Now $A(\xi)$ and $G(\xi)$ may be expressed in the following form

$$A(\xi) = \int_{-\infty}^{\infty} \alpha(\xi - u)\text{Ii}(u)du, \quad \dots\dots\dots (2)$$

$$G(\xi) = \int_{-\infty}^{\infty} \gamma(\xi - u)\text{Ii}(u)du, \quad \dots\dots\dots (3)$$

where $\text{Ii}(\xi) = \delta(\xi) + \frac{1}{2}\delta(\xi - \sigma)$, $\delta(\xi)$ is the unit impulse function, and σ is the separation of the two components of the $K\alpha$ doublet. We are here using the fact that

the intensities of the two components are in the ratio of 2 to 1. Solving these integral equations would yield $\alpha(\xi)$ and $\gamma(\xi)$, which are the forms that would be taken by $A(\xi)$ and $G(\xi)$ if the doublet were a single line. This is not a difficult step, but to clarify the exposition we shall defer the explanation to Appendix I, meanwhile assuming that $\alpha(\xi)$ and $\gamma(\xi)$ have been obtained from $A(\xi)$ and $G(\xi)$. Figure 2 shows $\alpha(\xi)$ and $\gamma(\xi)$.

TABLE I

$A(\xi) \times 10^3$	$F(\xi) \times 10^3$	$G(\xi) \times 10^3$
2	0	1
9	3	3
42	4	5
114	7	8
228	8	12
314	12	19
192	20	30
68	41	50
28	58	70
28	91	98
59	129	124
114	151	138
157	145	135
96	115	120
34	85	105
11	53	89
3	33	84
1	18	82
	12	80
	8	72
	6	58
	3	43
	0	28
		18
		11
		7
		4
		3
		2
		1

From equations (1), (2), and (3),

$$\int_{-\infty}^{\infty} \gamma(\xi - u) \text{Ii}(u) du = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(\xi - u) \alpha(u - w) \text{Ii}(w) du dw,$$

and from this it may be shown that

$$\gamma(\xi) = \int_{-\infty}^{\infty} \alpha(\xi - u) F(u) du, \quad \dots \dots \dots (4)$$

provided that the Fourier transform of $I\alpha(\xi)$ has no zeros, a property which may readily be verified. Equation (4) shows that the relationship between γ , α , and F is the same as that between G , A , and F as in equation (1).

The method of successive substitutions yields the solution of (4) as the following series :

$$F(\xi) = \gamma(\xi) + \varepsilon_1 + \varepsilon_2 + \dots,$$

where

$$\varepsilon_1 = \int_{-\infty}^{\infty} \gamma(\xi - u) [\delta(u) - \alpha(u)] du,$$

$$\varepsilon_2 = \int_{-\infty}^{\infty} \varepsilon_1(\xi - u) [\delta(u) - \alpha(u)] du, \text{ etc.}$$

This series converges if $|1 - \bar{\alpha}(s)| < 1$ for all s for which $\bar{\gamma}(s) \neq 0$, where $\bar{\alpha}(s)$ and $\bar{\gamma}(s)$ are the Fourier transforms of $\alpha(\xi)$ and $\gamma(\xi)$ (see Bracewell and Roberts 1954). When the condition is not met, the series may still give an asymptotic representation of $F(\xi)$, even though ultimately divergent (Bracewell, unpublished

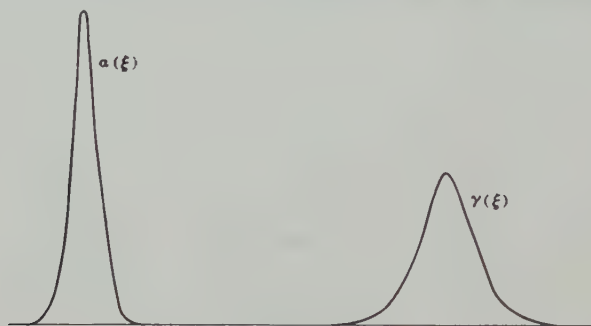


Fig. 2

data). It will be found that $A(\xi)$ fails to fulfil the convergence condition, on account of the violent disturbance due to the doublet. But in equation (4) this effect has been removed and, although technically the condition is still not met, it fails only for large values of s which do not play an important part if only a few stages of correction are required. The series is thus asymptotic to $F(\xi)$.

The working is shown in full in Table 2 for the present case. In the first column appear the values of α (Fig. 2), the instrumental profile corrected for doublet effect by the method of Appendix I. The column $\delta - \alpha$ is derived from the first column by subtracting the latter from the sequence $\dots 0 \ 0 \ 1 \ 0 \ 0 \dots$. It is written on a movable strip of paper and *in the reverse sense*. (The necessity for the reversal of sense will be found in the occurrence of the variable u with opposite signs in equations such as (1).) The column γ (Fig. 2) is obtained from the observed distribution by correction for doublet effect. The values of the first correction term ε_1 are then calculated by summing products of corresponding numbers in the second and third columns; e.g. the value marked with an asterisk is equal to $(-2) \times 1 + (-7) \times 3 + (-23) \times 5 + \dots$, and is written opposite the arrow. The remaining values are obtained by sliding the movable

strip up and down. Succeeding correction columns are obtained from their predecessors in exactly the same way.

Corrections up to 12 per cent. are introduced by ε_1 , and up to 3 per cent. by ε_2 . The next set of corrections, if calculated, will be found to be all less than 2 per cent., which, in the present case, is negligible. The last column, $\gamma + \varepsilon_1 + \varepsilon_2$, is thus the approximate solution yielded by the method of successive substitutions, and is shown as $F(\xi)$ in Figure 1.

TABLE 2

$\alpha \times 10^3$	$(\delta - \alpha) \times 10^3$ (read upwards)	$\gamma \times 10^3$	$\varepsilon_1 \times 10^3$	$\varepsilon_2 \times 10^3$	$\gamma + \varepsilon_1 + \varepsilon_2$
2	-2	1	-1	0	0
9	-7	3	-1	+1	3
42	-23	5	-2	+1	4
114	-67	8	-2	+1	7
228	-192	12	-4	0	8
314	+686	19	-6*	-1	12
192	-228	30	-8	-2	20
67	-114	49	-7	-1	41
23	-42	69	-7	-4	58
7	-9	95	-1	-3	91
2	-2	120	+8	+1	129
		132	+15	+4	151
		126	+15	+4	145
		105	+9	+1	115
		80	-5	0	85
		55	-0	-2	53
		36	-2	-1	33
		22	-3	-1	18
		14	-2	0	12
		9	-1	0	8
		6	-0	0	6
		3	-0	0	3
		1	-1	0	0

III. REFERENCES

- AVERBACH, B. L., and WARREN, B. E. (1949).—Interpretation of X-ray patterns of cold-worked metal. *J. Appl. Phys.* **20** : 885.
- BRACEWELL, R. N., and ROBERTS, J. A. (1954).—Aerial smoothing in radio astronomy. *Aust. J. Phys.* **7** : 615.
- DUMOND, J. W. M., and KIRKPATRICK, H. A. (1931).—Experimental evidence for electron velocities as the cause of Compton line breadth with the multicrystal spectrograph. *Phys. Rev.* **37** : 136.
- PATERSON, M. S. (1950).—Calculation of the correction for instrumental broadening in X-ray diffraction lines. *Proc. Phys. Soc. Lond. A* **63** : 477.
- PEASE, R. S. (1948).—The resolution of X-ray doublet diffraction lines into $\alpha_1\alpha_2$ components. *J. Sci. Instrum.* **25** : 353.
- RACHINGER, W. A. (1948).—A correction for the $\alpha_1\alpha_2$ doublet in the measurement of X-ray diffraction lines. *J. Sci. Instrum.* **25** : 254.

- STOKES, A. R. (1948).—A numerical Fourier-analysis method for the correction of widths and shapes of lines on X-ray powder photographs. *Proc. Phys. Soc. Lond.* **61** : 332.
- TAYLOR, A. (1941).—On the measurement of particle size by the X-ray method. *Phil. Mag.* **31** : 339.
- TAYLOR, A. (1945).—“An Introduction to X-ray Metallography.” (Chapman and Hall : London.)

APPENDIX I

Removal of Doublet Effect

To solve

$$A(\xi) = \int_{-\infty}^{\infty} \alpha(\xi - u) \text{II}(u) du$$

for $\alpha(\xi)$ when $A(\xi)$ is given and $\text{II}(\xi) = \delta(\xi) + \frac{1}{2}\delta(\xi - \sigma)$, we apply the method of successive substitutions.

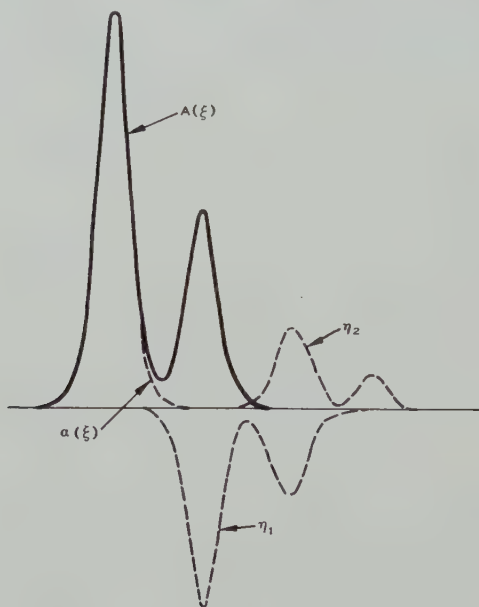


Fig. 3

Then

$$\alpha(\xi) = A(\xi) + \eta_1 + \eta_2 + \dots,$$

where

$$\eta_1 = \int_{-\infty}^{\infty} A(\xi - u) [\delta(u) - \text{II}(u)] du, \text{ etc.}$$

But this reduces to

$$\begin{aligned} \eta_1 &= -\frac{1}{2} \int_{-\infty}^{\infty} A(\xi - u) \delta(u - \sigma) du \\ &= -\frac{1}{2} A(\xi - \sigma), \end{aligned}$$

which is the same as $A(\xi)$ but changed in sign, halved, and shifted by σ . And η_2 must be obtained from η_1 in the same way. Hence finally

$$\alpha(\xi) = A(\xi) - \frac{1}{2}A(\xi - \sigma) + \frac{1}{4}A(\xi - 2\sigma) - \dots \quad (5)$$

Once this simple formula has been obtained, it is easy to understand (see Fig. 3). A sufficient condition that it should converge is that $|1 - \bar{\Pi}(s)| < 1$, where $\bar{\Pi}(s)$ is the Fourier transform of $\Pi(\xi)$, and this condition is met for all s .

In obtaining $\alpha(\xi)$ from $A(\xi)$ only one correction term need be taken, and in obtaining $\gamma(\xi)$ from $G(\xi)$ only two.

Equation (5) has been given by DuMond and Kirkpatrick (1931) who derived it by a direct argument. They state that the series converges, adducing the (insufficient) fact that the terms tend to zero. A process formally equivalent to equation (5), but adapted for greater convenience in use, has been given by Rachinger (1948),* and practical hints are contained in a note by Pease (1948). The present posing of the problem as an integral equation appears to be novel; it readily yields the proof of convergence, and shows that $\alpha(\xi)$ is not restricted to functions which fall to zero outside a finite range of ξ , nor even to functions which approach zero at all.

* The author is indebted to Dr. M. S. Paterson of the Research School of Physical Sciences, Australian National University, for the reference.

A GENERAL EQUATION FOR THE CHOICE OF GLASS FOR CEMENTED DOUBLETS

By W. H. STEEL*

[*Manuscript received June 21, 1954*]

Summary

A general equation is derived for the choice of glass for cemented doublet objectives for any object distance or residual aberration. The case where the aberrations are zero is studied in detail.

I. INTRODUCTION

In the design of a cemented doublet with prescribed amounts of longitudinal chromatic aberration, spherical aberration, and coma, two of the three variables required are provided by the radii of the surfaces, and the third must be chosen from the refractive indices and dispersions of the two glasses used. As the properties of available glasses cannot be represented by continuous variables, it is most economical of time if a suitable pair of glasses is chosen from those available as the first step of the design.

Brown and Smith (1946) have published tables which make possible the choice of glass when the object is at infinity and when the spherical aberration and coma are completely corrected. In this paper, a general equation is derived to cover the cases of finite object distances and uncorrected aberrations. The present analysis, like the work of Brown and Smith, is based on third order aberration theory and the results are strictly valid only for thin, low aperture objectives, but it is probable that the equation has wider application in practice, for the author has shown (Steel 1954) that glasses chosen from the Brown and Smith tables are suitable at all practical apertures.

The final equation obtained is a quintic in the relative dispersions of the two glasses. To avoid the need for solving such an equation, the special case where the aberrations are corrected is treated here in sufficient detail to cover the object distances used in practice. However, in the general case, this quintic would have to be solved, or an alternative approximate method given by Corrias (1953) could be used to find suitable glasses.

* Division of Physics, C.S.I.R.O., University Grounds, Sydney.

II. DERIVATION OF EQUATION

For a doublet of power K , the total curvatures c_a and c_b of the two components must satisfy the condition

$$(n_a - 1)c_a + (n_b - 1)c_b = K, \quad \dots \quad (1)$$

where n_a and n_b are the refractive indices of the two glasses used; in this paper they are taken for the e ray (5461 Å).

(a) Chromatic Aberration

If the chromatic aberration has a prescribed value, the two curvatures must also satisfy the condition

$$(n_a - 1)c_a/v_a + (n_b - 1)c_b/v_b = KA, \quad \dots \quad (2)$$

where v_a and v_b are the Abbe values for the two glasses, again taken here for the e ray, $v_e = (n_e - 1)/(n_F - n_C)$. The chromatic aberration is then given by

$$Lch'/l'^2 - Lch/l^2 = KA = 2\delta E/y^2, \quad \dots \quad (3)$$

l and l' being the object and image distance, y the semi-aperture, Lch and Lch' the longitudinal chromatic aberrations before and after the doublet, and δE the corresponding increase in the wave aberration ΣE .

Hence it follows that

$$c_a = \frac{\rho K}{(\rho - 1)(n_a - 1)}, \quad c_b = -\frac{K}{(\rho - 1)(n_b - 1)}, \quad \dots \quad (4)$$

where

$$\begin{aligned} \rho &= \frac{v_a(1 - A v_b)}{v_b(1 - A v_a)} \quad \dots \quad (5) \\ &= \frac{N + 1}{N - 1} \text{ in Brown and Smith's notation.} \end{aligned}$$

When the chromatic aberration is corrected, $\rho = v_a/v_b$.

(b) Coma

The third order coma may be represented by C where

$$C = 2OSC'/y^2 K^2 = 2\delta W_{\text{coma}}/y^3 e' K^2, \quad \dots \quad (6)$$

OSC' being the offence against the sine condition and δW_{coma} the increase in the coma wave aberration at an angular semi-field e' .

The standard equation for the third order coma of a cemented doublet (Conrady 1929) may now be solved for c_2 , the contact curvature, giving

$$\begin{aligned} c_2(\rho - 1)\{\rho(1 + a) - (1 + b)\} \\ = K[\{\rho + v(\rho - 1)\}\{\rho(2 + a) - (2 + b)\} - \frac{\rho^2}{1 - a} + \frac{1}{1 - b} + C(\rho - 1)^2], \dots \quad (7) \end{aligned}$$

where $a = 1/n_a$, $b = 1/n_b$, and $v = 1/Kl$, the reciprocal object distance at unit power.

The other curvatures are given by

$$c_1 = c_2 + c_a, \quad c_3 = c_2 - c_b, \quad \dots \quad (8)$$

(c) *Spherical Aberration*

The third order spherical aberration may be represented by S where

$$LA'_p/l'^2 - LA_p/l^2 = K^3 y^2 S/2 = 4\delta W_p/y^2, \quad \dots\dots\dots (9)$$

LA_p and LA'_p being the third order longitudinal spherical aberrations before and after the doublet and δW_p the corresponding increase in the wave aberration.

If this value S and the values of c_a , c_b , and c_2 are substituted in the standard equation for the primary spherical aberration of a doublet (Conrady 1929) the following equation is found connecting ρ and x , x being $\rho + v(\rho - 1)$,

$$\begin{aligned} & \{\rho(1+a) - (1+b)\}^2 \left[\frac{\rho^3}{(1-a)^2} - \frac{1}{(1-b)^2} - S(\rho-1)^3 + x\{\rho^2-1-4C(\rho-1)^2\} \right. \\ & \quad \left. - x^2\{\rho(5+2a) - (5+2b)\} \right] \\ & - \{\rho(1+a) - (1+b)\} \left[\frac{\rho^2}{1-a} - \frac{1}{1-b} - C(\rho-1)^2 - x\{\rho(2+a) - (2+b)\} \right] \left\{ \rho^2 \frac{2+a}{1-a} - \frac{2+b}{1-b} \right\} \\ & + \left[\frac{\rho^2}{1-a} - \frac{1}{1-b} - C(\rho-1)^2 - x\{\rho(2+a) - (2+b)\} \right]^2 \{\rho(1+2a) - (1+2b)\} = 0. \\ & \dots\dots\dots (10) \end{aligned}$$

For given values of the reciprocal object distance v , the aberrations C and S , and the refractive indices, this equation may be expanded as a quintic in ρ , giving a generalized form of Mossotti's equation (Argentieri 1954). Only positive roots are of practical interest; in the cases studied there is only one which is greater than unity for $n_a < n_b$ (the crown leading doublet) and less than unity for $n_a > n_b$ (the flint leading doublet). It should also be noted that the root for a given value of v is the reciprocal of the root of the equation when v is replaced by $-(v+1)$ and n_a and n_b interchanged, for this represents a reversal of the light path.

The Brown and Smith tables give solutions for the case $C=S=0$ when $v=0$ and when $v=-1$.

III. APPLICATION

(a) *General Case*

The general method of choosing a suitable pair of glasses would follow closely that used by the author (Steel 1954) for telescope objectives. One glass is chosen tentatively and the curve representing the glasses which can be used with it is plotted on a glass chart of n_e and v_e using points obtained by solving equation (10) for different indices of the unknown glass. If this curve does not pass sufficiently close to an available glass type, another initial choice would have to be made.

(b) *Corrected Doublets*

The numerical solution as a quintic is tedious but the reverse problem of finding the value of v for which a given glass pair is suitable is much simpler, involving only the solution of a quadratic. This method has been applied to the case of corrected doublets by solving for v for a typical selection of glass pairs.

When $C=S=0$, equation (10) may be expanded in powers of x to give

$$\begin{aligned} x^2(\rho-1)^2 - x \left[\{ \rho(1-a) - (1-b) \} \{ \rho(2+a) - (2+b) \} \left\{ \frac{\rho^2}{1-a} - \frac{1}{1-b} \right\} \right. \\ \left. - (\rho-1)(\rho^2-1) \{ \rho(1+a) - (1+b) \} \right] \\ + \{ \rho(2+a) - (2+b) \} \left\{ \frac{\rho^2}{1-a} - \frac{1}{1-b} \right\}^2 - \{ \rho(1+a) - (1+b) \}^2 \left\{ \frac{\rho^3}{(1-a)^2} - \frac{1}{(1-b)^2} \right\} \\ - (\rho^2-1) \{ \rho(1+a) - (1+b) \} \left\{ \frac{\rho^2}{1-a} - \frac{1}{1-b} \right\} = 0. \dots\dots\dots (11) \end{aligned}$$

The two roots of this equation coincide when

$$\left\{ (\rho-1)(\rho^2-1) + (\rho-b) \left(\frac{\rho^2}{1-a} - \frac{1}{1-b} \right) \right\}^2 = 4\rho(\rho-1)^2 \left(\frac{\rho}{1-a} - \frac{1}{1-b} \right)^2. \dots (12)$$

The value $\rho_{\min.}$ given by equation (12) is the minimum value of ρ corresponding to a given pair of indices n_a and n_b for which real solutions of v exist. The corresponding value of v is given by

$$v_{\min.} = -\frac{\rho}{\rho-1} + \frac{\frac{\rho^2}{1-a} - \frac{1}{1-b}}{(\rho-1)^2} - \sqrt{\rho \frac{\{ \rho(1+a) - (1+b) \} \left\{ \frac{\rho}{1-a} - \frac{1}{1-b} \right\}}{(\rho-1)^3}}. \dots (13)$$

Several common crown and flint glasses were selected from the catalogues of Messrs. Chance Bros. and of Messrs. Schott & Gen, and the results for doublets formed of these glasses is given in Table 1 in which are listed $\rho_{\min.}$ and $v_{\min.}$ and the two solutions of v , if real. Only the crown leading form is considered, as the flint leading form can always be considered in reverse.

TABLE 1
RECIPROCAL OBJECT DISTANCES FOR TYPICAL GLASS PAIRS

Glasses	Indices n_a	(e ray) n_b	v Value (e ray) a b		ρ v_a/v_b	Minimum Values		Solutions	
						$\rho_{\min.}$	$v_{\min.}$	v_1	v_2
BSC ₁ EDF	1.51160	1.65568	64.6	33.8	1.910	1.899	-1.13	-0.65	-1.58
	DF	1.62667	64.6	36.2	1.784	1.849	-1.05	—	—
BSC ₂ EDF	1.51951	1.65568	64.3	33.8	1.901	1.878	-1.11	-0.44	-1.72
	DF	1.51951	64.3	36.2	1.775	1.825	-1.02	—	—
HC EDF	1.52104	1.65568	60.6	33.8	1.792	1.874	-1.10	—	—
PSK3 EDF	1.55483	1.65568	63.5	33.8	1.876	1.778	-0.99	-0.44	-2.12
	F4	1.55483	63.5	36.8	1.726	1.684	-0.89	+0.03	-1.74
MBC EDF	1.57782	1.65568	57.9	33.8	1.712	1.712	-0.93	—	-0.93
	DF	1.57782	57.9	36.2	1.599	1.603	-0.83	—	—
DBC EDF	1.59113	1.65568	62.0	33.8	1.833	1.650	-0.88	+0.76	-2.27
	DF	1.59113	62.0	36.2	1.712	1.532	-0.78	+0.89	-2.27

The results are shown graphically in Figure 1 where $\rho/\rho_{\min.}$ is plotted against $v-v_{\min.}$ using extra results from the glass pairs BSC₁-EDF, PSK3-F4, and MBC-DF with hypothetical values of ρ . Although the points for the different glasses do not lie exactly on a common curve, they are sufficiently close to it within the range of v studied, for the author has shown (Steel 1954) that, for telescope objectives, the smallest tolerance on ρ is ± 0.01 .

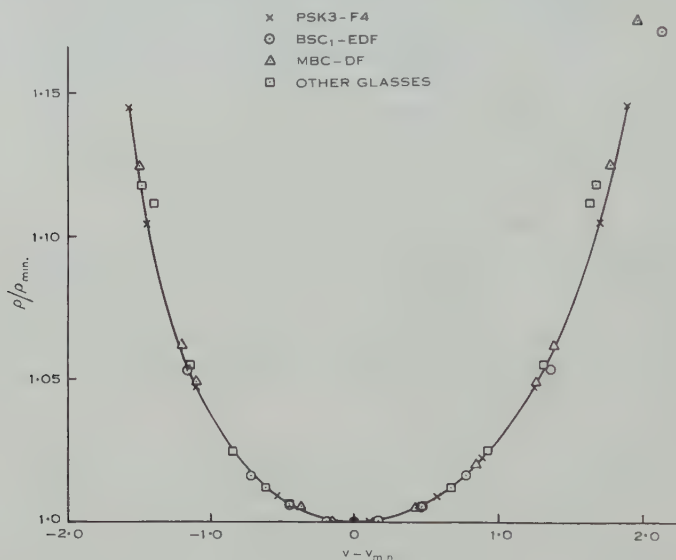


Fig. 1.—Variation of relative dispersion $\rho = v_a/v_b$ with reciprocal object distance v for unit power objectives.

This curve may be used to find the range of v that corresponds to a given tolerance on ρ . These latter tolerances are given in the article cited; for $\delta\rho = \pm 0.01$, Table 2 has been drawn up to show the appropriate glass pair for

TABLE 2
GLASSES SUITABLE FOR A GIVEN VALUE OF v WHEN $\delta\rho = \pm 0.01$

Range of v	-1.92, -1.85	-1.82, -1.64	-1.82, -1.58	-1.79, -1.72	-1.72, -1.34
Glasses ..	DF-DBC	PSK3-F4	BSC ₂ -EDF	EDF-DBC	BSC ₁ -EDF
Range of v	-1.35, -0.47	-1.19, -0.47	-1.12, -0.92	-0.93, -0.48	-0.72, -0.33
Glasses ..	MBC-EDF	MBC-DF*	F4-PSK3	BSC ₁ -EDF	EDF-PSK3
Range of v	-0.68, -0.41	-0.67, -0.28	-0.59, -0.32	-0.52, -0.07	-0.53, +0.19
Glasses ..	EDF-BSC ₂	PSK3-EDF	BSC ₂ -EDF	EDF-BSC ₁	DF-MBC*
Range of v	-0.53, +0.35	-0.08, +0.12	+0.34, +0.72	+0.58, +0.82	+0.64, +0.82
Glasses ..	EDF-MBC	PSK3-F4	EDF-BSC ₁	EDF-BSC ₂	F4-PSK3
Range of v	+0.72, +0.79	+0.85, +0.92	+1.07, +1.17		
Glasses ..	DBC-EDF	DBC-DF	EDF-PSK3		

* Although the pair MBC-DF has no real roots for v , the difference $\rho_{\min.} - v_a/v_b$ is 0.004 and so lies within the tolerance $\delta\rho = \pm 0.01$.

any given value of v in the range $-2 < v < 1$. Flint leading objectives are also included, the type of objective being shown by the order in which the glasses are listed.

It is seen that the range of real images, $-1 < v < 0$, is well covered by the glasses studied, and that the same glasses cover large regions of the total range of v . In particular, the four glasses BSC₁, BSC₂, MBC, and EDF can be used for all values of v from -1.82 to $+0.82$.

As the presence of chromatic aberration does not change the form of equations (11), (12), and (13), a similar method may be used for choosing the glass when this aberration is present, provided that ρ is derived from equation (5). But an extension of this method to the case of the other aberrations does not seem warranted and these cases would best be treated along the lines indicated in Section III (a).

IV. REFERENCES

- ARGENTIERI, D. (1954).—"Optica industriale." (Hoepli: Milan.)
BROWN, E. D., and SMITH, T. (1946).—*Phil. Trans. A* **240**: 59.
CONRADY, A. E. (1929).—"Applied Optics and Optical Design." (Oxford Univ. Press.)
CORRIAS, M. (1953).—*Atti Fond. Ronchi* **8**: 406.
STEEL, W. H. (1954).—*Aust. J. Phys.* **7**: 244.

THE DETERMINATION OF ELECTRON TRAJECTORIES IN THE PRESENCE OF SPACE CHARGE

By D. L. HOLLWAY*

[*Manuscript received July 15, 1954*]

Summary

The analogy between space charge in an electrostatic field and current sources in an electrolytic tank is considered and methods of simulating clouds of distributed charge by discrete current sources are developed.

The field patterns of charge regions of simple geometrical shapes are compared with those of sources and from these results rules are deduced giving the dimensions and arrangement of a source array suitable for any particular problem.

Experimental results show these principles applied to a plane diode, an axially symmetric electron beam, and a general asymmetric beam. It is concluded that in most problems space charge can be represented adequately in electrolytic models by a comparatively small number of sources.

I. INTRODUCTION

Although several different electron trajectory tracers have been described by Gabor (1937), Langmuir (1937, 1950), Marvaud (1948, 1952), and Sander, Oatley, and Yates (1949), comparatively few machines of this kind are in use. At the same time, problems requiring a knowledge of the paths of electrons or other particles arise frequently and are rarely in a form amenable to direct calculation.

One of the reasons why more use is not made of path tracers is that the effect of space charge is neglected. In many problems of practical interest this leads to appreciable errors. For example, a converging electron stream which is traced to a fine focus may be found to broaden excessively when the current is raised to a useful value, while deflected beams may be changed in position and shape by the influence of their charge. Frequently these effects limit performance, so that the tracings are least accurate when they would otherwise be of the greatest value.

In an attempt to overcome this limitation Musson-Genon (1947) devised an electrolytic tank in which the depth of the liquid was varied locally by an amount depending on the space charge density in that region. Recently Alma, Diemer, and Groendijk (1953) have solved space charge problems by adding distributed forces to models based on the rubber membrane analogy.

In the present work purely electrical methods have been developed, the space charge being represented by currents injected from sources spaced at intervals in an electrolytic model. Methods have been found of spreading the currents over an appreciable volume of the liquid in order to represent distributed charge and of finding the strengths and positions of the sources.

* Division of Electrotechnology, C.S.I.R.O., University Grounds, Sydney.

II. THEORETICAL BASIS

(a) *An Outline of the Method*

In the majority of space charge problems the charge distribution is unknown initially. As a first step, a family of trajectories is drawn for the given boundary conditions but neglecting space charge. A semi-automatic electron path tracer attached to an electrolytic tank is used, the paths being traced, without computation, in a few minutes. Sufficient paths are drawn to subdivide the beam into "tubes of flow" small enough to be considered to have uniform current density across any section. A pattern of current sources is then set up, using the principles to be described later, to correspond to the charge distribution predicted from the flow lines and the measured potential. The gradients from these sources modify the trajectories and the source currents are readjusted to suit the new flow lines. The successive approximations usually converge rapidly.

The use of current injection raises the problems of representing, by means of concentrated sources, charge distributed over an appreciable volume, as the sources are surrounded by strong local gradients which would cause distortion of the trajectories. It would be possible to reduce these by greatly increasing the number of current injection points, but as each of the currents has to be set individually, this becomes impossibly complicated. A solution which reduces to a minimum the number of injection points has been found from a consideration of the source field patterns and the results have been reduced to a number of simple rules for finding the dimensions, disposition, and strengths of the current sources.

In many problems it is only necessary to know whether a proposed electrode system is capable of producing a certain electron current density. By setting the current sources as described in Section II (f) this result is given by a single trial.

The procedure will now be considered in detail, beginning with the methods by which small concentrated current sources can be arranged to create field patterns in the model corresponding to those from distributed charges in the prototype.

(b) *The Distribution of the Injected Current*

When no space charge is present an electrolytic model may be used to determine the static distribution of potential U in an electrode system because the Laplace equation

$$\nabla^2 U = 0 \quad \dots\dots\dots (1)$$

is true also in a uniformly conducting medium free from sources. When charge is present, the fields are given by Poisson's equation*

$$\nabla^2 U = -\rho/\epsilon_0, \quad \dots\dots\dots (2)$$

where ρ =charge density, ϵ_0 =dielectric constant of free space.

* The rationalized M.K.S. system of units is used throughout this paper.

The corresponding expression for the model in an electrolyte of conductivity σ is

$$\nabla^2 U = -i'/\sigma, \quad \dots\dots\dots (3)$$

where i' may be termed the injected current density, although strictly speaking there is no physical means by which current may be injected in this way without disturbing the isotropy of the medium. However, the total charge δq appearing in the volume δV in the system may be represented in the model by the current δi where, from (2) and (3),

$$\delta i/\delta q = \sigma/\epsilon_0. \quad \dots\dots\dots (4)$$

The term σ/ϵ_0 is a conversion constant by which the current to be injected in the model may be found when the charge distribution in the original is known. The remaining problems are, firstly, that of distributing the current flow from a number of discrete sources in the model so that the gradient and potential patterns of the charge will be simulated as closely as possible and, secondly, of finding the charge distribution consistent with the boundary conditions.

In seeking methods of distributing the current, it seemed probable that the gradient and potential patterns at a liquid surface produced by current flow from a vertical line source, completely submerged some distance below, would be generally similar to those created by a uniform cylindrical column of charge in free space.

It was found that by choosing the dimensions of the source correctly the two fields could be matched closely.

To compare these fields consider a long cylinder of charge of uniform density ρ and radius r_1 . At points within the cylinder, $0 < r < r_1$,

$$dU/dr = -r\rho/2\epsilon_0, \quad \dots\dots\dots (5)$$

and the potential relative to that at some external radius r_2 is

$$U = (\rho r_1^2/4\epsilon_0)[1 - (r/r_1)^2 + 2 \ln r_2/r_1]. \quad \dots\dots\dots (6)$$

For radii outside the column

$$dU/dr = -r_1^2\rho/2\epsilon_0r, \quad \dots\dots\dots (7)$$

$$U = (\rho r_1^2/4\epsilon_0)(2 \ln r_2/r). \quad \dots\dots\dots (8)$$

These fields are to be reproduced by current from a thin source submerged as shown in Figure 1, in a layer of liquid bounded by a non-conducting floor. Let the current injected be iy_3 where y_3 is the depth of the liquid.

It may be noted immediately that at points remote from the source the current density is almost independent of depth, so that

$$dU/dr = -i/2\pi r\sigma, \quad \dots\dots\dots (9)$$

and the potential relative to that at radius r_2 is

$$U = (i/2\pi\sigma) \ln r_2/r. \quad \dots\dots\dots (10)$$

Thus, no matter how the current injection rod is changed in length or vertical position, the shapes of the remote gradient and potential curves match those

of the charge column and also agree in magnitude when in accordance with the conversion equation (4)

$$i = \pi r_1^2 \rho \sigma / \epsilon_0. \quad \dots\dots\dots (11)$$

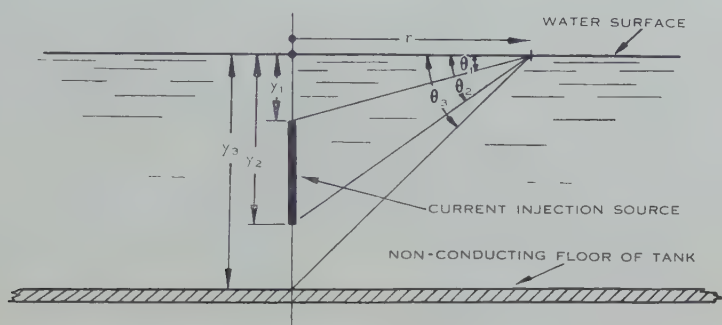


Fig. 1.—A current source used for representing space charge in a two-dimensional model.

To obtain the best match with the space charge field, the position of the gradient maximum must fall close to the radius r_1 . The field about a submerged wire has been calculated approximately and measured experimentally and it has been found that the peak occurs near the radius $r = y_1$, as shown in Figure 2.

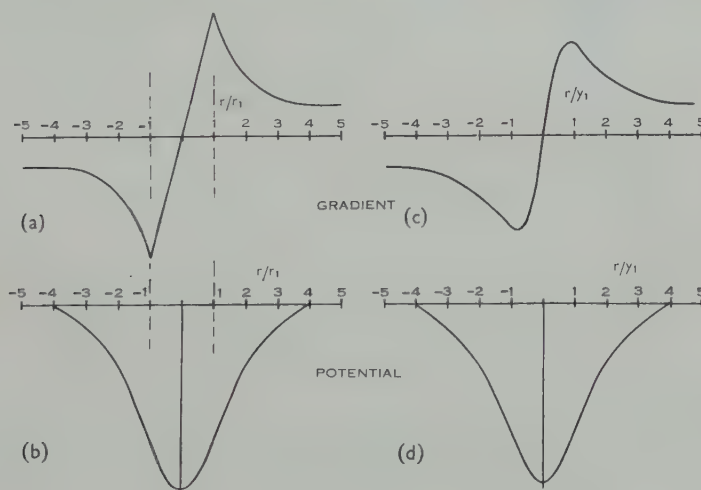


Fig. 2.—A comparison between the gradient (a) and potential (b) through a column of negative charge and the fields above a current source submerged as shown in Figure 1. The similarity between (a) and (c), (b) and (d) is used in representing space charge in two-dimensional models.

This result may be stated as a simple rule: in representing space charge columns by means of submerged line sources the depth of submersion must be equal to the radius of the column. This rule will be adhered to throughout this paper.

Although the position of the maximum is now correct, it is found that, when the source extends to the bottom of the tank ($y_2 = y_3$ in Fig. 1), the gradient of the source near its maximum is considerably smaller in magnitude than that of the charge column. Since the total current and y_1 are fixed, y_2 must be varied to improve the match and this can be done without disturbing the agreement at remote points. The determination of y_2 is considered in Appendix I. It is found that, when the line source is continuous between y_1 and y_2 , the depth y_2 should be equal to $0.5y_3$. The accuracy of match between the fields of a core of charge and current sources having $y_2/y_3 = 0.5$, but with different values of y_1/y_3 , is shown in Figure 3 by the curves (b), (c), and (d). When the ratio y_1/y_3 is less than 0.2 as in curve (d), the agreement can be improved by the method described in Appendix I, but this complication is rarely needed.

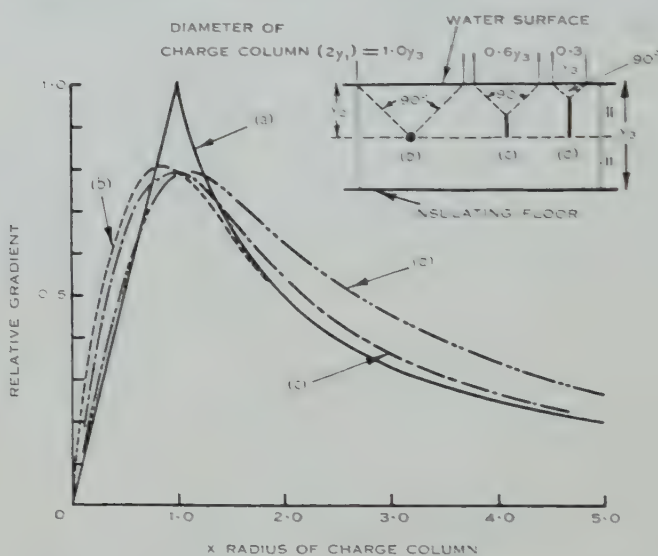


Fig. 3.—The gradient distributions about sources having different values of y_1/y_3 as shown in the key, compared with curve (a) for the equivalent charge column.

Thus the second rule for the representation of columns of space charge is that the depth of the bottom of the source should be made equal to half the depth of the tank.

We may now consider methods of grouping a number of simple sources to represent irregular and non-uniform regions of charge.

(c) The Representation of Layers and Charge Clouds of Irregular Shape

Most of the problems for which an electrolytic tank is useful either are two-dimensional or have axial symmetry and so may be represented in the wedge tank. The two-dimensional model will be considered first.

Suppose, for example, it is required to represent a vertical wall of charge bounded in the ZX -plane corresponding to the liquid surface as shown in

Figure 4 (a). To approximate to this layer, sources are arranged so that the corresponding columns of charge are overlapped to make the total area of the circles equal to the area of the layer. Curves 1 and 3 in Figure 4 (b) show that the gradient and potential measured along the Z -axis, with insulating sheets

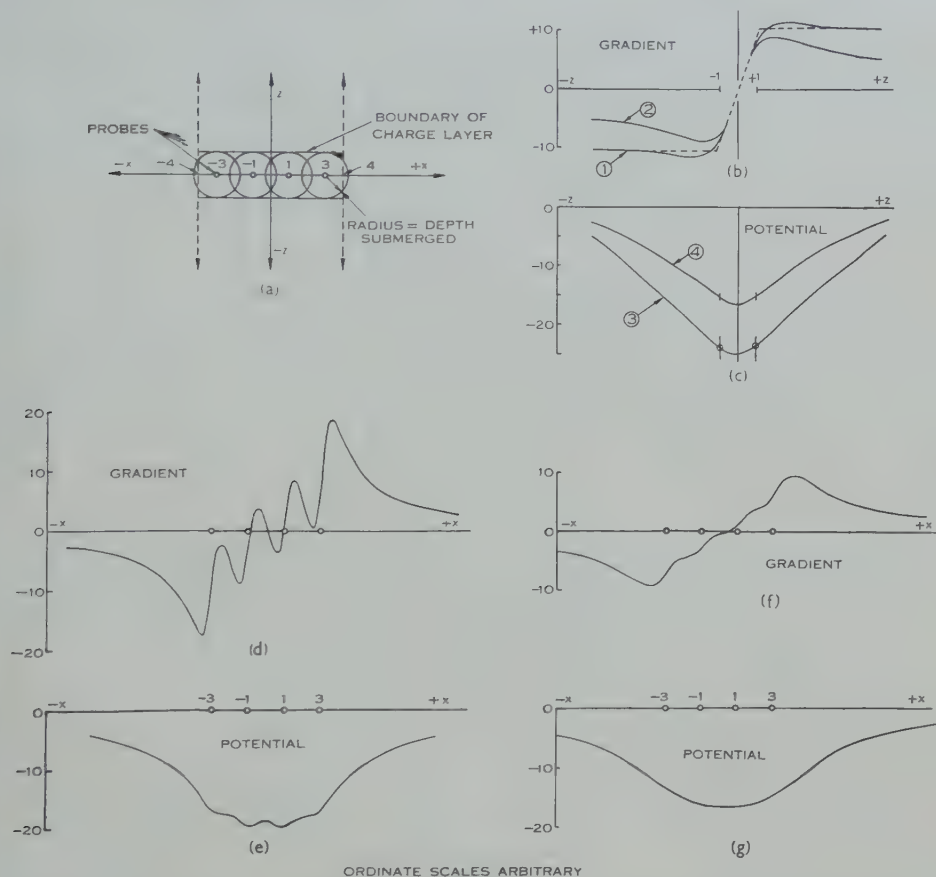


Fig. 4.—The method of combining current sources to simulate a layer of charge. (a) A view in the ZX -plane, the plane of the liquid surface, of four submerged sources and the equivalent layer of negative charge; (b) and (c) the gradient and potential along the Z -axis, curves 1 and 3 (only) were measured with insulating planes in the positions shown dotted in (a) and therefore refer to an infinite layer of charge having the gradient curve shown dotted in (b); (d) and (e) fields along the X -axis of the four sources shown in (a) submerged one-half the normal distance, $y_1 = r_1/2$; (f) and (g) as above, but with the sources fully submerged, $y_1 = r_1$. The local variations about each source have been greatly reduced.

in positions shown dotted in Figure 4 (a), are close to those of an infinite layer of charge. All the remaining curves in Figure 4 were measured with the insulating sheets removed and approximate to the fields of a finite layer of charge terminated at $z = \pm 4$ in Figure 4 (a). In combining sources in this way

advantage is taken of the fact that the field variations localized around each probe are attenuated very rapidly as the depth of submersion is increased. This is seen when Figures 4 (d) and (e), measured along the X -axis with the array submerged one-half the correct distance, are compared with the almost smooth curves (f) and (g) from sources at the correct depth, $y_1=r_1$.

This principle, of grouping probes to represent larger regions of charge, may be applied in the same way to charge clouds of irregular shape. The area representing the cloud in a two-dimensional tank is replaced by a pattern of overlapping circles of different sizes chosen to suit the boundary as accurately as necessary and having the same total area as the charge. The individual source currents are proportional to the charges represented.

(d) The Proximity Effect

When one or a line of sources is placed in close proximity to a conducting plane, the gradient at the liquid surface is slightly reduced by the presence of the plane. If the plane represents the surface of an emitting cathode where the magnitude of the gradient is important, it is necessary to correct for this reduction by increasing the source current. Figure 5 (a) shows a typical measurement

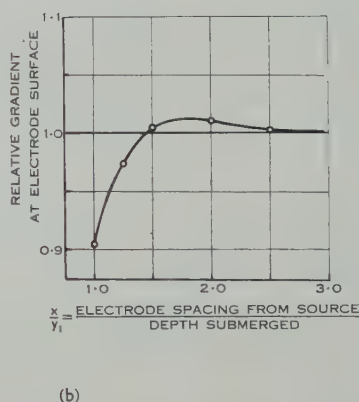
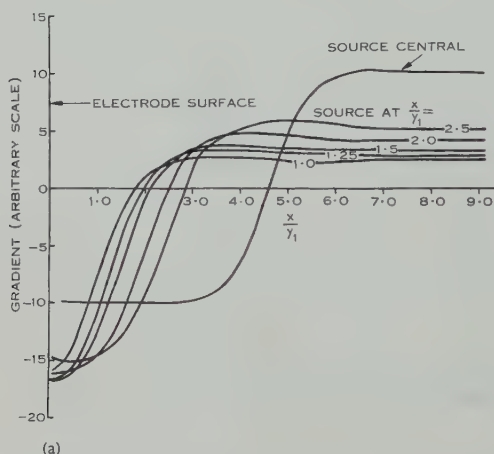


Fig. 5.—The proximity effect.

of this effect. A single constant-current source was placed in an insulated parallel-sided cell of width $\pi y_1/2$ closed by two perpendicular end electrodes and the gradient was measured for different positions of the source. As the change in gradient between the upper and lower levels of the curves in Figure 5 (a) should remain constant, the influence of the plane is shown by the lack of symmetry between the upper and lower crossings and in Figure 5 (b). At a distance $x=y_1$, the source current should be increased by 10 per cent. but at larger spacings no correction is needed. Although this method of correction raises the field on the other side of the source also, this position is usually less important than the cathode surface.

(e) *Charge in Axially Symmetric Systems*

Several different arrays have been used as sources in wedge tank models of axially symmetric systems. When the beam itself has axial symmetry and does not converge or diverge too rapidly, the charge may be represented by point sources placed along the axis as shown in Figure 6. Usually the spacing will be non-uniform. In order to reduce the number of points it is necessary to increase their spacing as much as possible. The spacing is limited by the need to keep local variations in the field at the edge of the beam small. Thus to find a suitable compromise it is necessary to consider the way in which these variations diminish away from the axis.

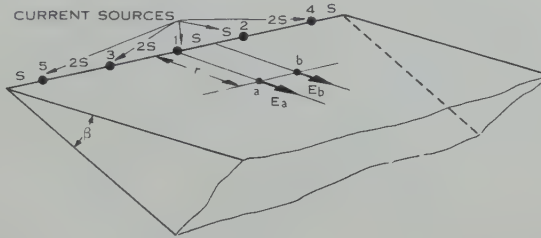


Fig. 6.—Point sources arranged along the axis of a wedge tank.

When an array is uniformly spaced, as shown in Figure 6, the maximum value of the radial gradient occurs at a and at corresponding points opposite each source. Thus, summing the contributions from an infinite line of sources each passing a current i ,

$$E_a = \frac{-i\alpha}{\beta\sigma S^2} \left\{ \frac{1}{2\alpha^3} + \sum_{N=1}^{N=\infty} \frac{1}{[\alpha^2 + (2N)^2]^{1.5}} \right\}, \quad \dots\dots\dots (12)$$

similarly, the minimum value E_b is given by

$$E_b = - \sum_{N=0}^{N=\infty} \frac{i\alpha}{\beta\sigma S^2} \frac{1}{[\alpha^2 + (2N+1)^2]^{1.5}}, \quad \dots\dots\dots (13)$$

where $\alpha = r/S$.

These expressions have been computed in the range $1 < \alpha < 2.2$ and checked experimentally. From the results shown in Figure 7 it is seen that the field of a row of point sources is equivalent to that of a line source having a superimposed space wave that varies in the axial direction and is attenuated rapidly away from the axis. The variation allowable along the trajectory depends on the conditions of the problem; for example α may have a smaller value where space charge is a correction rather than the controlling field. Usually it is safe to allow the trajectory to approach to $\alpha=2$, that is to within a radial distance equal to the spacing between sources, and in many cases α may be reduced to 1.0.

When the sources are placed along the axis, the trajectories being considered need not lie in the beam envelope. An inner path may be traced under space charge conditions by disregarding the shell of charge outside it and resetting the

source currents to correspond to the inner beam ; but it is necessary to check whether the resulting rise in the potential of the beam as a whole introduces an appreciable error.

The more general case of a beam travelling away from the axis will now be considered. It is assumed that the beam electrode system can be represented to the desired accuracy either in a normal wedge tank or in one in which the wedge angle is π , and that the centre line of the beam always lies in the plane of the water surface. The problem then is to represent the half beam charge using sources as sparingly as possible. When the envelope is known approximately, the choice falls on submerged point sources arranged as shown in

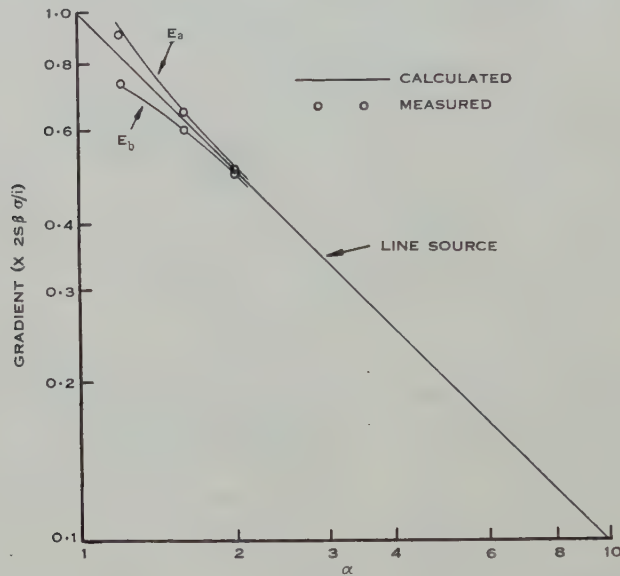


Fig. 7.—The maximum and minimum values of the radial field created by the sources shown in Figure 6.

Figure 8, the depth of submergence being made equal to the radius of the electron stream r_1 in order to bring the position of the gradient maximum to the edge of the stream and to reduce local field variations within the envelope. As the probe cannot approach closer than distance r_1 to the sources, these may be considered equivalent to a line of strength $i/\delta l$ and an image above the liquid surface. The wedge angle is assumed to be sufficiently large to allow other images to be neglected. The mean gradient at the edge ab is then

$$dU/dr = -i/2\pi r \sigma \delta l, \dots\dots\dots (14)$$

i.e. one-half the value obtained from sources placed at the central positions c', d', e', \dots . Thus the source current calculated from the beam space charge

$$i/\delta l = \pi r^2 \rho \sigma / 2 \epsilon_0 \dots\dots\dots (15)$$

cannot be used directly but must be increased to twice this value. The edge field is then correct and that within the beam nearly so. The remote field is

too large, but in many problems this is less important. The variation in the edge field at positions a and b (Fig. 8) is less than that shown in Figure 7 for $\alpha = \sqrt{2}$ because the angle between the source and image gradients is less at the mid positions such as b . The variation measured experimentally was ± 4 per cent.

When the beam envelope is unknown initially it is simpler to place the line of sources near the water surface and set to the currents indicated by equation (4). If the sources are at the centre of the beam the gradient is correct for the edge electrons, whatever changes occur in the radius of the envelope, but paths within the beam cannot be traced using the same source currents.

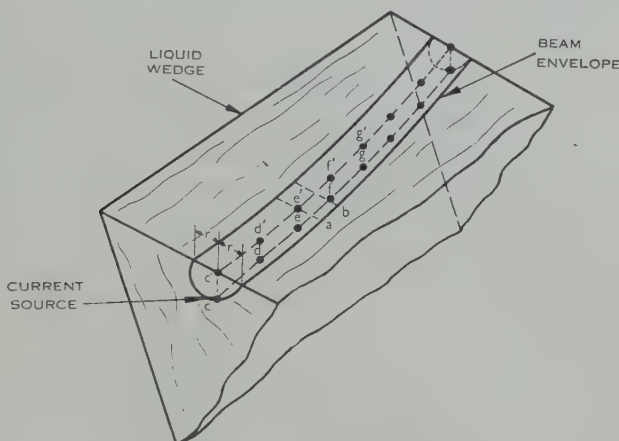


Fig. 8.—An array of sources c, d, e, f, \dots representing charge in a beam moving off the axis in a wedge model. The alternative positions c', d', e', f', \dots are used in tracing the trajectories defining the beam envelope.

When the paths converge to a focus of small radius, it is usually inconvenient to increase the number of sources sufficiently to define it exactly. In these cases it is simpler to calculate the minimum radius from the convergence and radius at some earlier position, using the space charge formulae referred to in Section III (b) and ignoring fields other than that of the beam charge.

(f) *Scaling and the Determination of Charge Density*

The electrolytic model will usually differ from the original electrode structure both in size and in the scale of the applied potentials. In representing space charge, the potential scale, which may be ignored in solving the Laplace equation, must be taken into account in determining the values of the source currents. When the electron paths are influenced by space charge, it may be shown that the paths in the model will correspond exactly with those in the original if the relationship between the current and potential in the model is such that the perveance remains the same as in the original system.

As a first step, electron paths are drawn for the space-charge-free field. These paths form boundaries of "tubes" of electron current flow. Assuming

that the current in a given tube in the original is known to be I_0 , the current in the corresponding tube of the model I_m is

$$I_m = I_0 (U_m/U_0)^{1.5}, \quad \dots\dots\dots (16)$$

where U_0 , U_m are potentials in the original and model respectively.

If U'_m , U'_0 are the mean potentials in a small length δl of the current tube which is to be represented in the tank by one source, the total charge δq , in this length δl is

$$\delta q = \frac{\delta I_m}{(2e/m)^{0.5} U'_m{}^{0.5}} = \frac{\delta I_0 (U_m/U_0)^{1.5}}{(2e/m)^{0.5} U'_m{}^{0.5}}, \quad \dots\dots\dots (17)$$

and the source current i is given by

$$i = \frac{\delta q \sigma}{\epsilon_0} = I_0 \frac{\delta l U_m^{1.5} \sigma}{U'_m{}^{0.5} U_0 (2e/m)^{0.5} \epsilon_0}. \quad \dots\dots\dots (18)$$

(In M.K.S. units the numerical value of $1/(2e/m)^{0.5} \epsilon_0$ is 1.9053×10^5 .)

When sources are being set to represent axial beams in the wedge tank, the separate measurement of the conductivity σ may be avoided by measuring the gradient at one or more values of the radius r , where, if β is the wedge angle,

$$i/\delta l \beta r = dU/dr. \quad \dots\dots\dots (19)$$

Therefore from (18) and (19)

$$\left[\frac{\beta r}{U_m} \cdot \frac{dU}{dr} \right]_r = \frac{I_0/U_0^{3/2}}{(U'_m/U_m)^{0.5} (2e/m)^{0.5} \epsilon_0}. \quad \dots\dots\dots (20)$$

When the current I_0 is not known initially, as for example in structures which include a cathode surface, the source currents are set so that for all sources

$$i U'_m{}^{0.5} / \delta l = K. \quad \dots\dots\dots (21)$$

The value of K is then increased until the off-cathode gradient falls to zero. The value of I_0 then follows from (18). For similar electrode structures the perveance is independent of size, so that the real dimensions of the original electrodes are needed only for the calculation of the cathode current density, transit time, and so on, from the model measurements.

The potential U'_m has been defined as the mean potential in the interval. When the ratio of the potentials at the beginning (U_1) and the end (U_2) of the interval differs appreciably from unity, the source current can be set more accurately by taking as the value of U'_m a potential corresponding to the mean velocity of an electron in the interval. Thus

$$U'_m = \left[(x_1 - x_2) / \int_{x_1}^{x_2} \frac{dx}{\sqrt{U}} \right]^2. \quad \dots\dots\dots (22)$$

The curves in Figure 9 were calculated from (22) and the relation

$$U \propto x^N,$$

where $N=1.3$ applies to a plane diode and the exponents 1.2 and 1.5 are, within the limits indicated, approximations to the results of Langmuir and Blodgett (1924) for space charge limited flow between spherical surfaces. Thus

the curve in Figure 9 marked $N=1.2$ refers to a comparatively rapidly expanding tube of flow and that marked $N=1.5$ refers to one contracting. The small difference between these indicates that in setting currents near the cathode surface values from curve $N=1.3$ may be used in most cases and elsewhere, when $U_1/U_2 > 0.3$, the mean potential may be taken for U'_m .

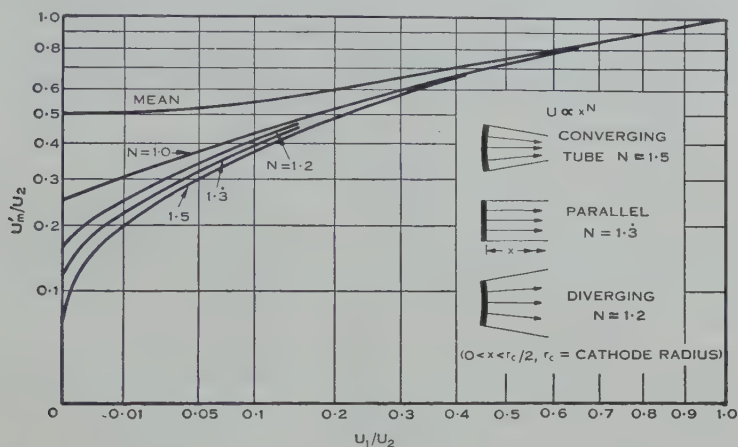


Fig. 9.—Values of U'_m in terms of U_1 and U_2 for converging, parallel and diverging space charge limited flow. In the cases $N \approx 1.5$ and ≈ 1.2 , $x = |r_c - r|$.

III. EXPERIMENTAL RESULTS

(a) The Plane Diode

Three examples will be given to illustrate the use of these methods in space charge problems. The results of two of these may be compared with analytical solutions so that the accuracy likely to be achieved in similar problems may be estimated.

The first is a model of an infinite parallel plane diode consisting of an insulated cell having electrodes and current sources arranged as shown in Figure 10. The sources are sleeves of 0.1 mm brass supported on 3 mm glass tubes and connected through individual variable resistances to a common point. The potential U'' of this point is adjustable between zero and 110 V, 50 c/s. When, as in this model, a tube of flow terminates on an emitting cathode, the gradient at the cathode surface is much more sensitive to nearby sources than to those further away. This offers a more rapid method of finding the source currents than that based on the Laplace distribution. First the cathode gradient is reduced to zero by adjusting the first two sources alone to a common value of K (eqn. (21)). All the remaining sources are then set directly to this value and the potential U'' is reduced until the cathode gradient returns to zero. A sufficiently accurate solution usually follows one further relaxation of the source currents.

The diode model was solved in this way and the results are compared with the theoretical distribution in Figure 10. The perveance, calculated from the

mean value of K and a separate measurement of the conductivity was within 1 per cent. and the individual values were within ± 5 per cent. of the correct result.

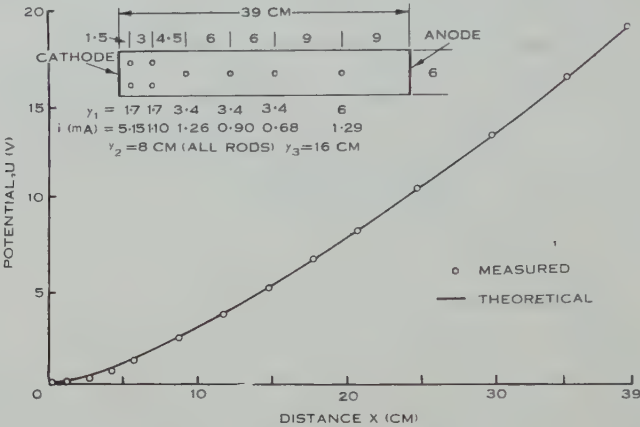


Fig. 10.—The solution for a space charge limited, infinite parallel plane diode using six sources in an insulated cell. The currents shown for the first two sources are each divided equally between two rods in order to fulfil the requirement that $y_1=r_1$.

(b) Defocusing of a Convergent Beam

The second example shows the effect of space-charge in expanding the focus of a uniformly convergent electron beam. In this test, values of the product $k\tau$, where k is a function of the perveance given by

$$k^2=33 \times 10^{-6} U^{1.5}/I, \dots\dots\dots (23)$$

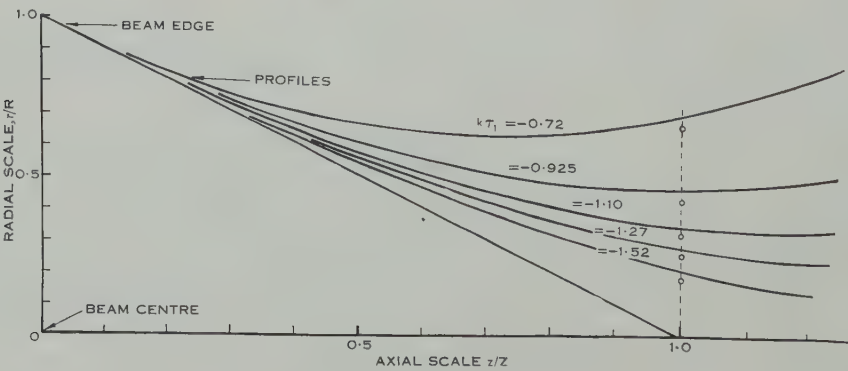


Fig. 11.—The envelope trajectories of a convergent axially symmetric beam defocused by space charge.

- Copied from experimental tracings.
- ooooooo Calculated assuming $Z \gg R$ (see Section III (b)).

and τ is the initial convergence R/Z (Fig. 11), were chosen to correspond to those already published (Hollway 1952). The source currents were set in accordance with equation (20) by the direct measurement of the gradient as described in Section II (f). The presence of the sources results in a variation in potential

along the axis of the beam which is ignored in the calculation. By rearranging the circuit of the electron path tracer paths were drawn for both constant and varying beam potential and the difference was shown to be small. As the shapes of the calculated paths agree closely with the tracings shown in Figure 11, for clarity only a single calculated point is given for each path. While part of the difference shown is experimental error, the greater part is explained by the fact that the calculation is approximate for such rapidly converging beams because the axial electron velocity, instead of the total velocity, is assumed to be equivalent to the beam potential, whereas the experimental solution is more accurate.

(c) A Deflected Beam

The third is an example of the more general problem of predicting the shape of a beam moving off the axis of an axially symmetric system. In the absence of space charge the initially parallel trajectories leaving electrode *A* in Figure 12 (a) are brought to a focus some distance in front of the anode *B* by

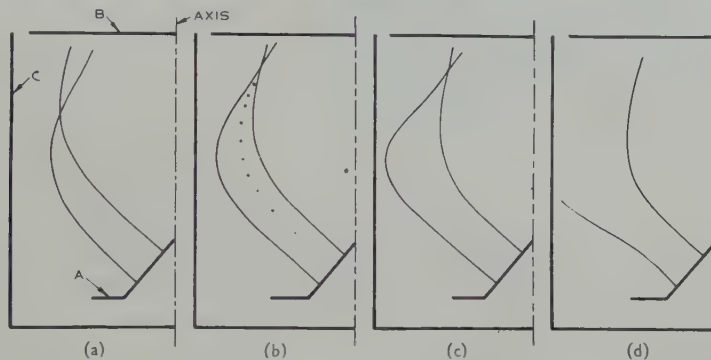


Fig. 12.—A general space charge problem showing the change in form of a deflected beam with increasing current. Electrode potentials: *A* 250 V, *B* 500 V, *C* zero. (a) Calculated beam current $I_0 = 0$, (b) $I_0 = 0.8$ mA, (c) $I_0 = 1.0$ mA, (d) $I_0 = 1.8$ mA.

the field of a cylinder *C* at zero potential. Sources were set at unequal intervals along the centre of the beam, slightly submerged to miss the moving probe of the trajectory tracer. The spacings were made less than the local beam diameter as shown in Figure 12 (b). After measuring the potentials at the beginnings and ends of the intervals and setting the source currents, usually only one revision was necessary to obtain a sufficiently constant value of *K*. The paths drawn in Figures 12 (b), (c), and (d) show the expansion of the beam with increasing beam current. It may be predicted from the change in curvature of the paths approaching the electrode *C* that the overall focus would be broad, especially in (c). When the current is raised still further, much of the beam is dispersed.

IV. CONCLUSIONS

It is shown that by arranging current sources according to the rules given, the influence of space charge can be examined quantitatively in most of the usual electrolytic models whether or not the electron current is known initially.

Although the convergence of the source currents to their final values is quite rapid, it would be a useful extension of the technique to include in each source lead a control circuit designed to keep the current proportional to $1/\sqrt{U_m}$. The electron current in a given tube of flow could then be set directly and the bounding trajectories found more rapidly.

V. ACKNOWLEDGMENT

The author wishes to thank Professor D. M. Myers for his interest and for making facilities available in the Department of Electrical Engineering, Sydney University, where the present work was carried out.

VI. REFERENCES

- ALMA, G. A., DIEMER, G., and GROENDIJK, H. (1953).—*Philips Tech. Rev.* **14**: 336-44.
 GABOR, D. (1937).—*Nature* **139**: 373.
 HOLLWAY, D. L. (1952).—*Aust. J. Sci. Res. A* **5**: 430.
 LANGMUIR, D. B. (1937).—*Nature* **139**: 1066.
 LANGMUIR, D. B. (1950).—*R.C.A. Rev.* **11**: 143-54.
 LANGMUIR, I., and BLODGETT, K. B. (1924).—*Phys. Rev.* **24**: 49-59.
 MARVAUD, J. (1948).—*C. R. Acad. Sci., Paris* **226**: 476-8.
 MARVAUD, J. (1952).—*C. R. Acad. Sci., Paris* **234**: 45-7.
 MUSSON-GENON, R. (1947).—*Ann. Télécomm.* **2**: 298-320.
 SANDER, K. F., OATLEY, C. W., and YATES, J. G. (1949).—*Nature* **163**: 403.

APPENDIX I

The Optimum Length of a Vertical Line Source

As described in Section II (b), a cylindrical column of space charge may be represented by a vertical line source extending from the depth y_1 to y_2 in the layer of liquid shown in Figure 1. To find the optimum value of y_2/y_3 an expression must be derived for the potential gradient at the liquid surface. For the present purpose it may be assumed that at all positions along the wire the current per unit length flowing into the liquid is constant and equal to

$$iy_3/(y_2 - y_1).$$

The gradient in the plane of the liquid surface is found by assuming an image of the source above the surface and, when r is not small compared with y_3 , images below the floor and above the floor-image must be included also. It is not necessary to specify these exactly, however. They may be replaced by uniform line sources extending from $+$ and $-y_3$ to infinity. The gradient then becomes

$$\frac{dU}{dr} = \frac{-ir}{2\pi\sigma} \left\{ \frac{y_3}{y_2 - y_1} \int_{y_1}^{y_2} \frac{dy}{(y^2 + r^2)^{1.5}} + \int_{y_3}^{\infty} \frac{dy}{(y^2 + r^2)^{1.5}} \right\} \dots\dots (24)$$

$$= \frac{-i}{2\pi\sigma r} \left\{ \frac{y_3}{y_2 - y_1} (\sin \theta_2 - \sin \theta_1) + (1 - \sin \theta_3) \right\}. \dots (25)$$

In order to match the field just outside the core of charge, (25) must become identical with (9). The bracketed term of (25) is therefore set equal to unity and y_2/y_3 is found as a function of r/y_1 for different values of the parameter

y_2/y_1 . As y_2/y_3 is reduced, the field in the region of the maximum is increased, raising the curves (b), (c), and (d) in Figure 3 with respect to (a). As the crests of these curves are less sharp than (a), a compromise is needed to obtain the best general match and the value $y_2/y_3=0.5$ has been chosen. The curves (b)-(d) show the comparatively small effect of variations in y_1 when y_2/y_3 is kept constant and equal to 0.5.

When $y_1/y_3 < 0.2$, as in (d), a somewhat closer agreement can be obtained by using, instead of a line source between y_1 and y_2 , one extending from y_1 to y_3 but including equally spaced gaps. Experiment shows that a suitable length for the conductors is $0.5y_1$ and for the gaps $2.5y_1$. Thus the sources extend from y_1 to $1.5y_1$, $4y_1$ to $4.5y_1$, and so on, until y_3 is reached.

THE EFFECT OF MULTIPLE SCATTERING ON ELECTRON ENERGY LOSS DISTRIBUTIONS

By D. F. HEBBARD* and P. R. WILSON*

[*Manuscript received November 29, 1954*]

Summary

Calculations have been performed to obtain the theoretical energy loss distributions of 1 MeV electrons traversing foils thick enough for multiple scattering to be important. A criterion has been developed and applied to previous experiments to determine the foil thicknesses for which multiple scattering has an appreciable effect on the energy loss distribution.

I. INTRODUCTION

The problem of the distribution of energy losses in the passage of high energy electrons through foils has recently been studied by several workers (Paul and Reich 1950 ; Chen and Warshaw 1951 ; Kalil and Birkhoff 1953 ; West 1953 ; Hungerford and Birkhoff 1954). Under various conditions of foil thickness and electron energy, these workers have found values for the most probable energy loss which agree reasonably well with the theories of Landau (1944) and of Blunck and Leisegang (1950). The Blunck-Leisegang distribution is, however, appreciably wider than that of Landau for thin foils, but the correction to the Landau distribution becomes less significant as the foil thickness is increased whereas the difference between the experimental and theoretical widths increases with thickness. Most explanations of this difference have been qualitative statements that multiple scattering would increase the theoretical widths. As a step towards clearing the confusion that exists, it has been thought desirable to determine the extent to which multiple scattering affects the energy loss distribution of electrons and to apply these predictions to previous experiments. In conjunction with this, an experiment has been performed under conditions similar to those used in the calculations. A comparison will be made with this experiment in an accompanying paper (McDonell, Hanson, and Wilson 1955).

II. THE MONTE CARLO METHOD

(a) *General Procedure*

The C.S.I.R.O. Mark I Digital Computer was used to determine the paths of particles obeying one of two multiple scattering theories : that of Rossi and Greisen (1941) adapted to the three-dimensional case, and that of Molière (1948). The Molière theory should be more accurate, so most of the calculations were based on this, but two cases were worked out using the Rossi-Greisen theory

* Physics Department, University of Melbourne.

for comparison (see Table 2 and Figs. 1 and 2). These multiple scattering theories are applied to short sections of fixed length l along the path of the particle. In each section the particle is considered to go straight until it reaches the end of the section when the direction of travel and the energy are changed for the next section. Each particle is incident normally on the scattering foil with an energy of 1 MeV and is followed through the foil until the energy falls below 0.4 MeV or until the particle escapes from the foil. Particles whose energy falls below 0.4 MeV are considered lost as the scattering at that energy is becoming excessive. Also classified as lost are those particles leaving the source side of the foil.

In common with other methods, the Monte Carlo method requires an accurate knowledge of the whole angular distribution for scattering. The small angle approximation used in multiple scattering theory means that the probability of large deflections, though small, is not necessarily accurate. A large deflection usually means a greatly increased path length and thus a large energy loss. The Monte Carlo distribution should therefore be accurate for small and medium energy losses but not necessarily accurate in predicting the probability of large energy losses or the stopping of electrons in the foil.

(b) Choice of Energy Loss Distribution

The energy loss distribution chosen was that of Landau because it is more convenient mathematically. The angular distribution of scattered electrons depends on their energy, but this dependence is slight and there is little change in the shape of the angular distribution over the range of energies predicted by the Landau theory for electrons with the same path length in the foil. This statement will only be true for electrons with kinetic energies above 0.5 MeV because, below this energy, the shape of the angular distribution begins to change rapidly with energy. With this reservation, the Landau and Blunck-Leisegang theories will agree in the prediction of the angular distribution of scattered electrons and of the increase in path length, because of the agreement in their prediction of the most probable energy loss.

(c) Choice of Section Length

The choice of the section length l was governed by the following considerations:

(i) there must be a sufficient number of individual collisions to make valid the equations on which the energy loss and multiple scattering theories are based;

(ii) the r.m.s. angle of scattering must not be too large since the multiple scattering theories involve a small angle approximation;

(iii) the mean path length in each section must be close to the section length so that the multiple scattering and energy loss theories may be applied in each section without modification.

The values of section length l finally chosen are set out in Table 1. For heavy elements it is clearly impossible to reconcile the requirement (i) with (ii) and (iii) above.

(d) Results

The distributions obtained from the Monte Carlo calculations are shown in Figures 1 and 2 (full curves). The method of obtaining smooth curves was to plot the histograms with an interval of 0.01 MeV. Where these points did not

TABLE I
BEHAVIOUR OF A 1 MEV ELECTRON IN CERTAIN ABSORBERS

Element	Al	Au
Section length, l (mg/cm ²) ..	5.26	3.68
Average number of collisions (Molière 1948)	44	34
ξ/I_0Z (Landau 1944)	2.47	0.24
R.M.S. angle of scattering (from Molière function f°)	5.3°	9.7°
Mean path length/foil thickness (Yang 1951)	1.01	1.04
Most probable energy loss (Landau) .. (keV) ..	4.1	1.8

lie on a smooth curve, larger intervals were chosen and a curve of best fit drawn. The position of the most probable loss was determined by finding the locus of the mean of the abscissae of the two branches of the smoothed curve at equal ordinates and extrapolating to the maximum height. The widths and peak

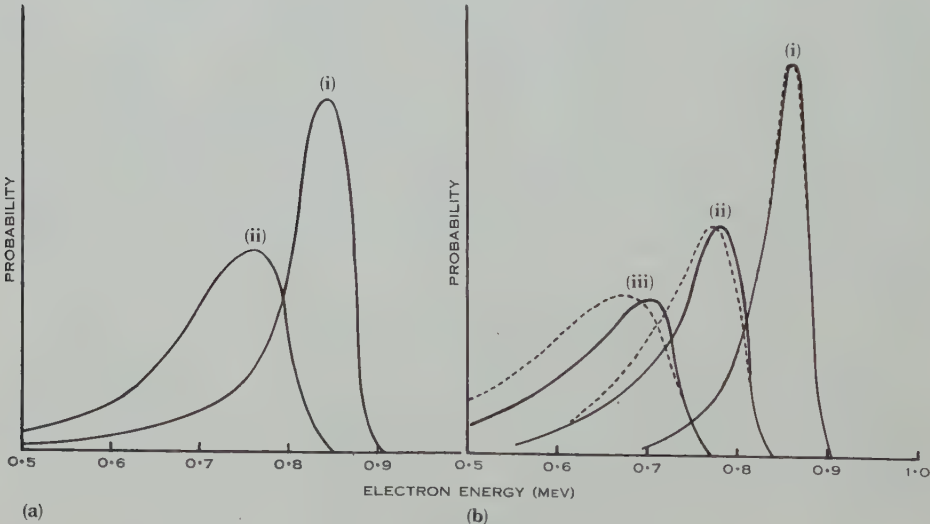


Fig. 1.—Energy distributions of 1 MeV electrons after passing through aluminium; (i) 105 mg/cm², (ii) 158 mg/cm², (iii) 210 mg/cm². (a) Monte Carlo (Rossi-Greisen theory) distribution. (b) Full curve: Monte Carlo (Molière theory) distribution. Dashed curve: folded (Landau, modified Yang) distribution.

positions of these distributions and the percentage of particles lost are compared in Table 2 with the Landau distribution, and are plotted against absorber thickness in Figure 3.

III. NUMERICAL INTEGRATION METHOD

The effect of multiple scattering on the energy loss distribution has also been calculated by folding the Landau distribution with a path length distribu-

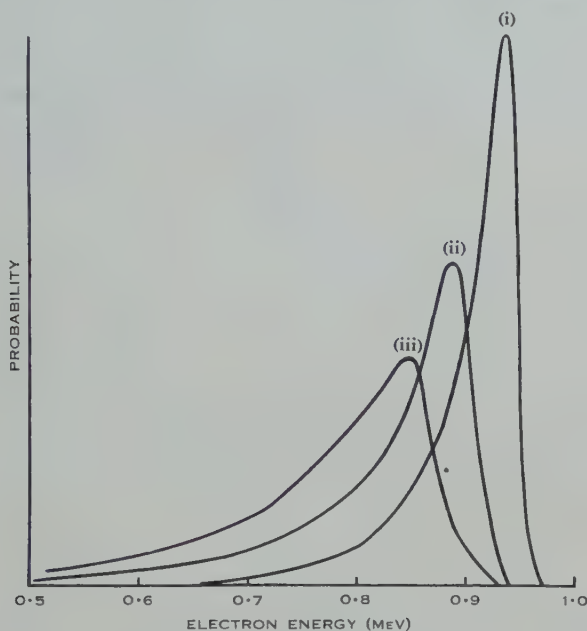


Fig. 2.—Energy distribution of 1 MeV electrons after passing through gold: Monte Carlo (Molière theory); (i) 73 mg/cm², (ii) 110 mg/cm², (iii) 147 mg/cm².

tion. Yang (1951) gives the only available path length distribution, based on the Rossi-Greisen multiple scattering theory. It was found necessary to modify Yang's distribution to give better agreement with the Monte Carlo (Molière

TABLE 2

COMPARISON OF THE LANDAU THEORY WITH THE MONTE CARLO CALCULATIONS FOR 1 MEV ELECTRONS AT NORMAL INCIDENCE

Element and Thickness (mg/cm ²)	Most Probable Loss (keV)			Half Height Width (keV)			Loss of Electrons (%) Monte Carlo*
	Monte Carlo*	Monte Carlo†	Landau	Monte Carlo*	Monte Carlo†	Landau	
Al 105	140	155	125	54	75	37	0.2
„ 158	218	236	194	89	134	55	1.5
„ 210	295		264	129		74	8.4
Au 73	63		52	45		21	2.7
„ 110	111		81	75		32	5.0
„ 147	154		113	114		43	8.2

* Monte Carlo: 1000 particles obeying the Molière scattering theory.

† Monte Carlo: 500 particles obeying the Rossi-Greisen scattering theory.

theory) path length distribution. The mean increase of path length according to Yang is nearly twice the mean increase calculated from the theory of Goudsmit and Saunderson (1940) as corrected by Rose (1940). The Goudsmit-Saunderson theory uses a better small angle approximation than the Rossi-Greisen theory. Yang's theory was therefore modified by multiplying his predicted increase of path length (case I) by the factor $\frac{1}{2}$, and the resultant distribution is compared in Figure 4 with the Monte Carlo and the unmodified Yang path length distributions

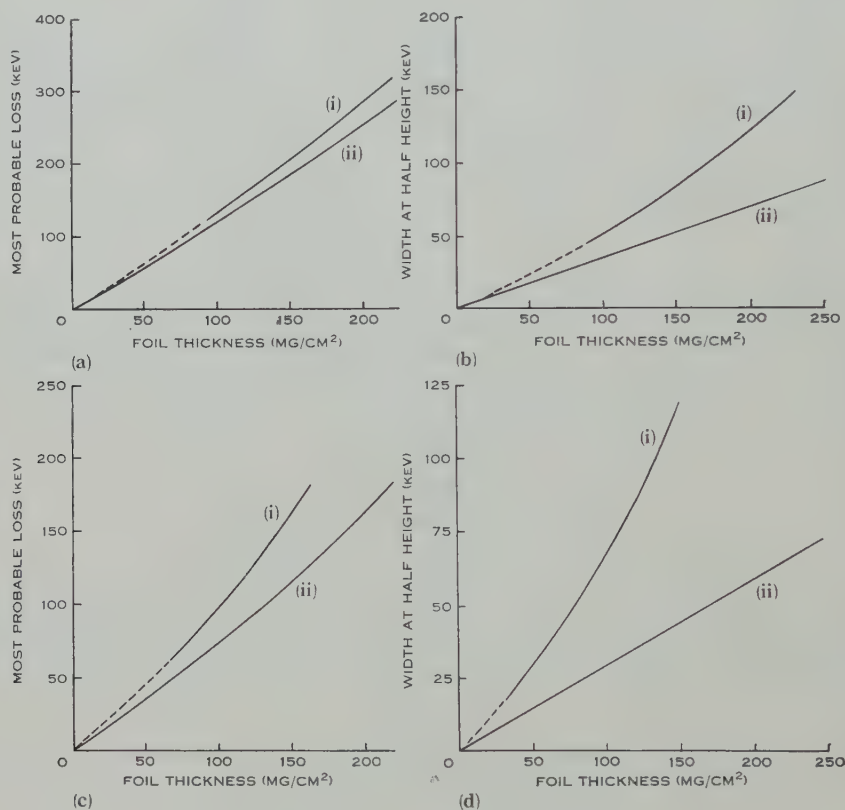


Fig. 3.—1 MeV electrons incident normally on foils; (i) Monte Carlo (Molière theory), (ii) Landau theory. (a) Aluminium: most probable energy loss. (b) Aluminium: width at half height. (c) Gold: most probable energy loss. (d) Gold: width at half height.

for a thick aluminium foil. Similar agreement is obtained between the Monte Carlo and modified Yang distributions for a thick gold foil and even better agreement for thinner foils.

The factor $\frac{1}{2}$, giving the correct mean increase of path length, should be valid over a wide range of elements, foil thicknesses, and electron energies, because, although the dependence on these quantities is different for Yang's theory when compared with that of Goudsmit and Saunderson, the difference is in a logarithmic term which varies only slowly.

Figure 1 (b) (dashed curves) gives the result of folding the modified Yang path length distributions with the appropriate Landau distributions for 1 MeV electrons incident normally on aluminium, while Table 3 gives a comparison

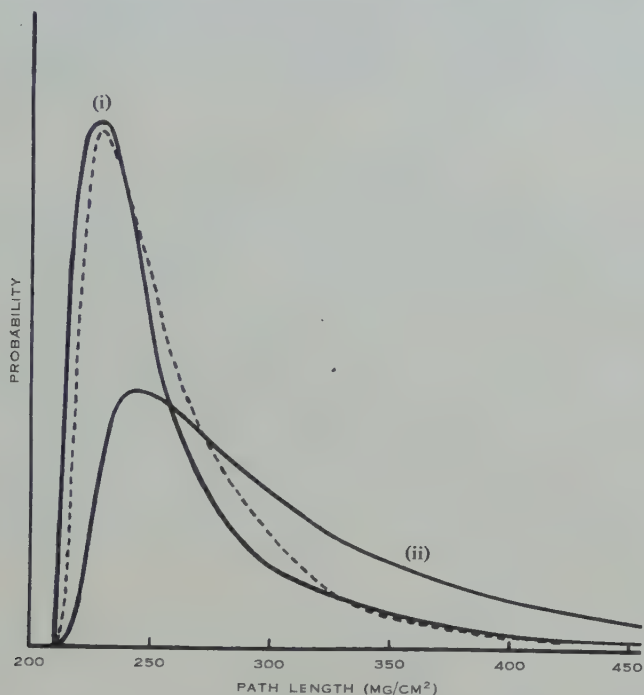


Fig. 4.—Path length distribution of 1 MeV electrons incident normally on 210 mg/cm² of aluminium. (i) Full curve: Monte Carlo (Molière theory) distribution. Dashed curve: modified Yang distribution. (ii) Unmodified Yang distribution.

of widths at half height and most probable losses for the Monte Carlo and folded distributions, and also the estimate of most probable loss obtained by multiplying the Landau value by the ratio of mean path length (modified Yang theory) to

TABLE 3
1 MeV ELECTRONS INCIDENT NORMALLY ON ALUMINIUM

Thickness (mg/cm ²)	Most Probable Loss (keV)			Half Height Width (keV)	
	Monte Carlo*	Folded†	Landau $\times \bar{s}/t^\ddagger$	Monte Carlo*	Folded†
105	140	140	140	54	54
158	218	228	229	89	101
210	295	322	328	129	170

* Monte Carlo: 1000 particles obeying the Molière scattering theory.

† Folded: the Landau folded with the modified Yang distribution.

‡ \bar{s}/t is the ratio of the mean path length to foil thickness.

foil thickness. It can be shown theoretically that the first effect of multiple scattering is to increase the width at half height and the most probable energy loss by this ratio. Application of this criterion to the Monte Carlo calculations, where peak positions and widths have an error of the order of 3 KeV, leads to thicknesses of 10 sections for both aluminium and gold as the minimum thickness for which multiple scattering is important under these conditions. The agreement with Figure 3 is reasonable.

IV. APPLICATION TO EXPERIMENT

The distributions obtained from these calculations are compared directly with the experiment described in the following paper (McDonell, Hanson, and Wilson 1955) and are further discussed there. However, it is of some interest to apply the criteria developed in these calculations to some experiments to determine the part played by multiple scattering.

Although Chen and Warshaw (1951) claim agreement between experimental and theoretical most probable loss, multiple scattering is important in all cases except the thin beryllium and aluminium foils. Thus the agreement for heavier elements must be somewhat fortuitous. On the other hand multiple scattering is not important in the experiments of West (1953), using various inert gases as absorbers. He finds most probable losses which are on the average 90 per cent. of those predicted by Landau and Blunck-Leisegang, and widths intermediate between these two theories.

In the experiments of Hungerford and Birkhoff (1954) which repeated, more accurately, the work of Kalil and Birkhoff (1953), multiple scattering would be important for all except the thin aluminium and copper foils, if the beam were incident normally. However, the geometry of the source-foil arrangement is such that incident particles are accepted in an angular range from 45 to 55° although only particles emergent at 45° are accepted by the spectrometer. Because of this, multiple scattering will cause significant path length increase even for the thin aluminium and copper foils.

V. CONCLUSIONS

Two conclusions can be drawn from the work at this stage. Firstly, the Monte Carlo and the folded distributions for aluminium foils of medium thickness correspond very closely. Therefore, in further calculations of this type, the use of the path length integration method would obviate the considerable labour involved in the Monte Carlo method but retain the same validity. From Figure 1, the energy distribution is sensitive to the multiple scattering theory assumed, and therefore to the assumed path length distribution in the numerical integration method.

Secondly, it appears that the previous workers have not made sufficient allowance for the effect of multiple scattering when comparing experiment with theory. The agreement claimed by these workers for foils of heavier elements indicates that the Landau (or Blunck-Leisegang) most probable losses and the Blunck-Leisegang widths at half height are actually too large, and this is supported by the work of West in cases where multiple scattering is not important.

VI. ACKNOWLEDGMENTS

The authors wish to thank Professor L. H. Martin and Associate Professor C. B. O. Mohr for their interest in the work ; Mr. K. C. Hines and Mr. J. A. McDonell for many helpful discussions ; and the staff of the C.S.I.R.O. Digital Computer, Radiophysics Division, Sydney (Mr. T. Pearcey and Mr. G. W. Hill) for their assistance in the calculations.

VII. REFERENCES

- BLUNCK, O., and LEISEGANG, S. (1950).—*Z. Phys.* **128** : 500.
CHEN, J. J. L., and WARSHAW, S. D. (1951).—*Phys. Rev.* **84** : 355.
GOUDSMIT, S. A., and SAUNDERSON, J. L. (1940).—*Phys. Rev.* **57** : 24.
HUNGERFORD, E. T., and BIRKHOFF, R. D. (1954).—*Phys. Rev.* **95** : 6.
KALIL, F., and BIRKHOFF, R. D. (1953).—*Phys. Rev.* **91** : 505.
LANDAU, L. (1944).—*J. Phys., Moscow* **8** : 201.
McDONELL, J. A., HANSON, M. A., and WILSON, P. R. (1955).—*Aust. J. Phys.* **8** : 98.
MOLIÈRE, G. (1948).—*Z. Naturf. A* **3** : 78.
PAUL, W., and REICH, H. (1950).—*Z. Phys.* **127** : 429.
ROSE, M. E. (1940).—*Phys. Rev.* **58** : 90.
ROSSI, B., and GREISEN, K. (1941).—*Rev. Mod. Phys.* **13** : 240.
WEST, D. (1953).—*Proc. Phys. Soc. Lond. A* **66** : 306.
YANG, C. N. (1951).—*Phys. Rev.* **84** : 599.

THE ENERGY LOSS DISTRIBUTIONS OF 1 MeV ELECTRONS IN ALUMINIUM FOILS

By J. A. McDONELL,* M. A. HANSON,* and P. R. WILSON*

[*Manuscript received November 29, 1954*]

Summary

The energy loss distributions for 1 MeV electrons traversing aluminium foils of various thicknesses have been investigated experimentally. Several of the foils were of thicknesses such that the distributions were expected to show considerable effects due to multiple scattering. It was found that the most probable energy losses were in good agreement with those predicted by the calculations of Hebbard and Wilson (1955).

I. INTRODUCTION

The failure of the theories of Landau (1944) and Blunck and Leisegang (1950) to give an accurate description of the energy loss distributions for electrons has been discussed in the preceding paper (Hebbard and Wilson 1955), hereafter referred to as I, in which the effect of multiple scattering on the path length in the foil has been considered. Two theoretical distributions were obtained for foils of thickness 0.105, 0.157, and 0.210 g/cm². The first set were derived from a "Monte Carlo" calculation of energy loss and the second by numerically folding the Landau distribution for a particular foil, with a path length distribution obtained from Yang's calculation (1951). Better calculations by Goudsmit and Saunderson (1940) indicated that the mean increase in path length was half that calculated by Yang, and in the folding process the shape of Yang's distribution was retained but the path length increases were reduced by half. For brevity, the results of these calculations will be referred to as the Monte Carlo and the Yang distributions.

The most desirable geometry for experimental investigation of energy loss distributions is that in which a narrow, parallel beam of electrons falls normally on an absorber, and those emerging are detected by a spectrometer whose solid angle of acceptance is known. The two most interesting cases would be those in which the solid angle is either very small, so that only particles passing "straight through" are detected, or approximately 2π , in which case all emergent particles are detected. In either case, a reasonably simple comparison with theory is possible only if the incident beam is parallel.

If the shape of the energy loss distributions is to be measured in detail, apparatus of considerable precision is required, since the distributions (in the region of validity of the Landau theory) are quite narrow; the width at half height is of the order of a few per cent. of the particle energy at most. This

* Physics Department, University of Melbourne.

means that the energy spread of the incident beam and the resolving power of the spectrometer should both be at least an order of magnitude smaller. If they are not, one has to resort to folding techniques to compare experiment with theory, with a consequent loss of accuracy in the comparison. On the other hand, a good deal more latitude is possible if the aim is primarily an investigation of the most probable energy loss. The spread of energy of the incident beam is not nearly so important for energies in the relativistic region, because the most probable energy loss depends on the incident energy only through the term $(v/c)^2$.

II. APPARATUS

A monochromatic beam of electrons was selected by a short magnetic lens β -ray spectrometer of a type similar to that described by Deutsch, Elliot, and Evans (1944). Because of its pure β -spectrum, ^{32}P , of 20 mc activity, was used as the electron source, giving a beam of the order of 11,000 electrons per minute. The electron trajectory was defined by baffles shown in Figure 1, and ultimately

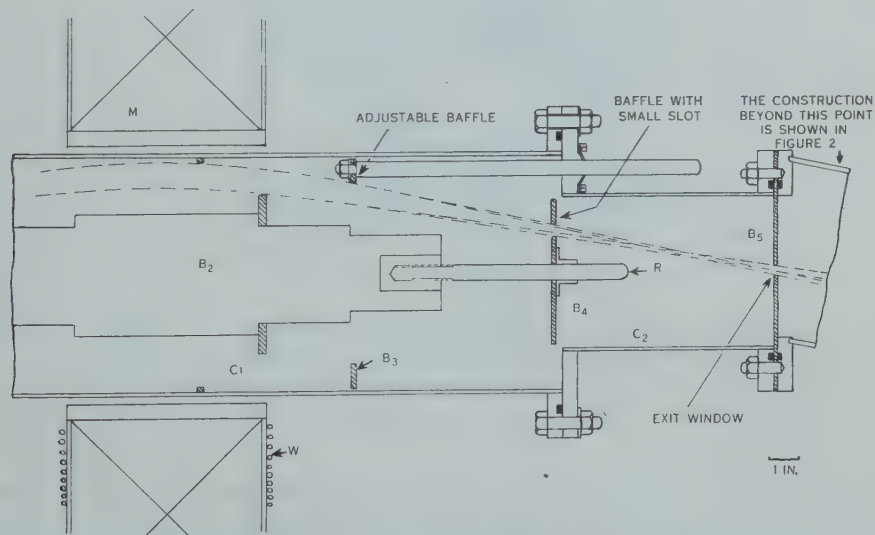


Fig. 1.—Baffle system producing narrow emergent beam.

a rectilinear beam of electrons travelling at an angle of 11° to the axis of the spectrometer was selected by baffles B_4 and B_5 . The only opening in B_4 was an eccentric slot whose radial width was slightly wider than the spread of the trajectories at this point, and which subtended an angle of 15° at the spectrometer axis. The baffle B_5 was situated at the focus, and contained a $\frac{1}{4}$ in. diameter circular aperture at its centre. The transmission of the spectrometer was thus reduced by a factor of $1/24$, but the remaining trajectories passing through the aperture in B_5 had an angular divergence of less than 1.5° . Although the resultant transmission of this arrangement was low (about 8×10^{-5}) a beam of reasonable intensity could be obtained, since the spectrometer was acting only as an energy selector, and so it was not necessary to use a thin source.

The energy spread of the beam was estimated at 2.3 per cent. at half-intensity, by observing the shape of the conversion line from the 364 keV γ -ray of ^{131}I , and allowing for line broadening due to the thickness of the source.

Figure 2 shows the way in which various absorbers were placed in front of the beam. A short "adaptor" section enabled the foil holder to be set at an angle of 11° to the spectrometer axis, so that the beam fell normally on the foils. The foil holder was an aluminium disk, with six equally spaced holes of 1 in. diameter. A bevel gear and a driveshaft passing through an O-ring seal enabled the foil holder to be rotated from outside the vacuum system. Foils were mounted over four holes in the holder, the fifth hole being left blank, and the sixth blocked off with an absorber thick enough to stop the electrons completely.

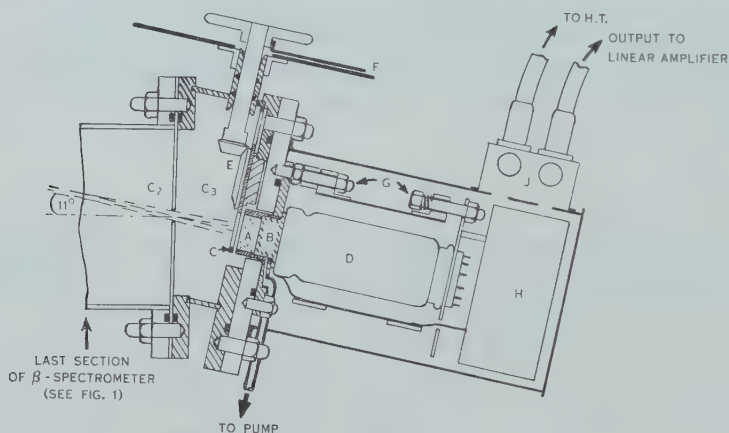


Fig. 2.—"Adaptor" section and scintillation counter. *A*, NaI(Tl) crystal; *B*, glass light guide; *C*, aluminium foil holder; *D*, photomultiplier; *E*, bevel gears; *F*, pointer indicating position of foil holder; *G*, spring-loaded mounting for photomultiplier; *H*, potential divider; *J*, preamplifier.

A scintillation counter was used as a spectrometer-detector. In order to obtain maximum counting rates, and also to obtain a greater effect due to scattering, the crystal was mounted as shown in Figure 2 so that its face was 0.10 in. from the plane of the foil holder. To obtain the best resolution with this type of detector, an NaI(Tl) crystal was used. The crystal was 1 in. in diameter and 0.5 in. long; this length was excessive for the purpose of stopping the electrons, but was used because it was rather easier to handle. Since the source of the electrons was ^{32}P , backgrounds were at all times negligible, so that the size of the crystal was not a disadvantage from that point of view.

The mounting of the crystal presented some difficulties. To obtain good resolution, the crystal must be surrounded by an efficient reflector; but, on the other hand, the amount of additional absorbing material between the crystal and the foils had to be kept as small as possible. Therefore the reflector next to the plane surface of the crystal had to be a very thin aluminium foil. However, the whole crystal chamber had to be airtight—because of the hygroscopic nature

of sodium iodide—and at the same time capable of being placed inside a vacuum system. The arrangement eventually evolved is shown in Figure 3. The reflector surrounding the crystal consisted of a highly polished aluminium cylinder, with an aluminium foil (3.5 mg/cm^2) cemented across its end. The crystal was attached to a glass block with silicone grease. The crystal and glass block were then slid into the cylinder (which was made a neat fit to the crystal) and the system was sealed with an O-ring between the flange of the glass block and the brass plate on which the whole arrangement was mounted. The glass block was held secure by a brass ring, overlapping the flange, and held to the brass plate by six screws which could be accurately adjusted so that the crystal

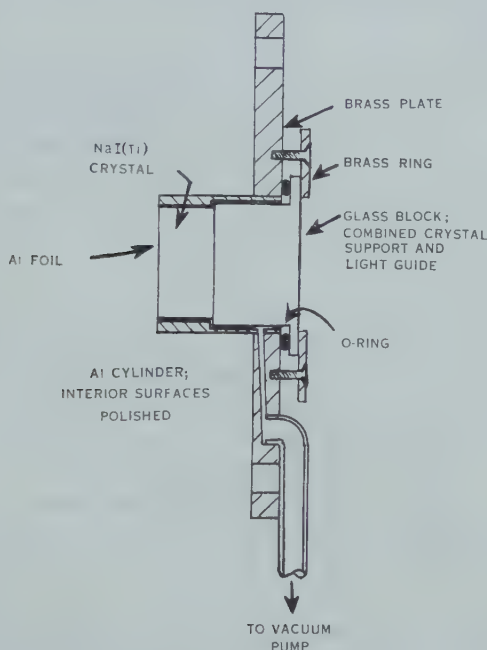


Fig. 3.—Crystal chamber.

penetrated the desired depth into the cylinder. With this adjustment there was very little flexure of the thin aluminium foil as it flattened against the crystal face when the chamber was evacuated. This small vacuum system was independent of that of the spectrometer, and was continuously pumped to a pressure of 10^{-3} mm Hg .

The crystal surfaces were left rough ground as recommended by Borkowski and Clark (1953), except for the face next to the glass block, which was polished. The glass block served as a light guide to the photocathode of an EMI 6260 photomultiplier. It was possible either to remove the photomultiplier and its mounting while leaving the crystal assembly in the spectrometer, or to remove the whole scintillation counter as a unit. The pulses, after amplification, were analysed by a single channel analyser.

III. EXPERIMENTAL PROCEDURE

(a) Adjustment of the Spectrometer

After adjusting the alignment of the tube and the position of the slotted baffle by the usual criterion, viz. obtaining a maximum count rate for electrons of a given energy, the spread of the beam in the plane of the absorbers was investigated photographically. Its cross section was found to be quite well defined, and of roughly oval shape, with major and minor axes of approximately 0.35 and 0.25 in. Thus there was no danger of scattering from the edges of the holes over which the foils were mounted and, when the beam was central, all electrons scattered through angles less than 73° and most of those scattered through less than 82° should be detected by the crystal. The spectrometer calibration was obtained using a ^{131}I source.

(b) Performance of the Scintillation Counter

It was found, from comparison of the peak in the differential pulse height distribution for 1 MeV electrons and the peaks produced by irradiation of the crystal with γ -rays from ^{131}I and ^{137}Cs , that, for the particular high tension across the photomultiplier required to produce optimum resolution, the relation between the energy of electrons striking the scintillator and the amplified pulse height from the photomultiplier was non-linear.

It was found that with a low H.T. the ratio of the pulse heights for the ^{137}Cs and the 1 MeV electron peaks was 0.57, whereas we expected a ratio of 0.664.* Furthermore, this ratio increased almost linearly with increasing H.T., and passed through the value 0.664. This effect was at first attributed to some possible peculiarity of behaviour of the photomultiplier in a magnetic field, and the H.T. was adjusted to that value which gave the ratio of 0.664, in order to secure a linear relation between pulse height and energy. Later, at various lower currents in the magnet coil, the pulse heights of the ^{137}Cs peak and of the peaks for electrons of the energy focused in the spectrometer were measured. These were plotted as a function of the energy of the focused electrons, and it was immediately seen that these curves did *not* cross at an energy of 662 keV, as was expected. This meant that the pulse height produced by electrons of energy E incident on the flat face of the crystal was not equal to the pulse height produced by γ -rays of energy E . The difference is presumably due to the different spatial distributions of the scintillations produced in the crystal in the two cases.

It was thus necessary to plot a calibration curve of pulse height against incident electron energy. This was made difficult by the fact that the amplification of the photomultiplier varied with the magnetic field produced by the current in the field coils of the spectrometer. To measure the pulse height of electrons selected from the spectrometer at a particular energy, with the amplification which obtains for 1 MeV electrons, the electron beam was blocked off and the pulse height corresponding to the 661.65 keV γ -ray from ^{137}Cs was

* $(\text{Energy of Cs } \gamma\text{-ray})/(\text{Energy of incident electrons}) = 0.662/0.997 = 0.664$. The use of 0.997 MeV instead of 1 MeV for the incident electron energy is due to an estimated energy loss of 0.003 MeV in the thin aluminium foil covering the crystal face.

determined at the current for 1 MeV electrons. Then, with the magnet current set at the energy under consideration, the photomultiplier H.T. was varied until the original pulse height for the ^{137}Cs γ -ray was obtained. With this H.T. the pulse height for the focused electrons was determined and this procedure was repeated for several values of the electron energy.

(c) *The Energy Loss Runs*

An incident energy of 1 MeV was used throughout. Two sets of four absorbers were used, their thicknesses being 0.010, 0.026, 0.035, 0.052, 0.083, 0.101, 0.153, and 0.215 g/cm². In each run, the pulse height distributions for the four absorbers, the blank hole, and the blocked one were obtained, three runs being done for each of the two sets of absorbers. In each run the count rates for each of the six foil holder positions were taken in turn at every bias setting of the pulse height analyser, so that the readings on all the peaks were taken at about the same time. In this way it was hoped to minimize any effects due to undetected drifts in the circuitry.

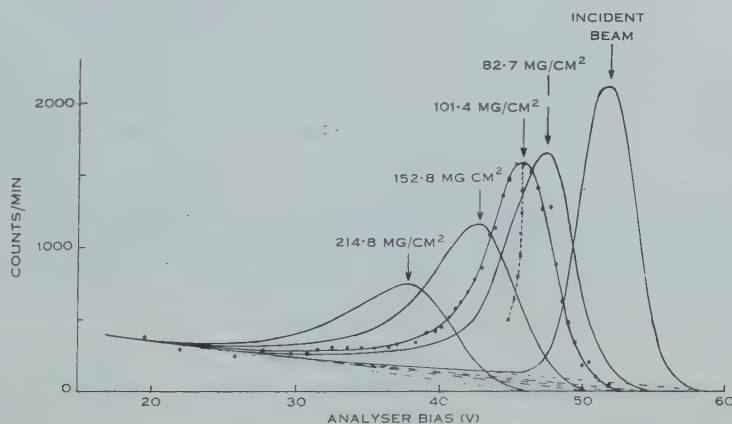


Fig. 4.—Results of one experimental run. The scatter of the points about the estimated smooth curve is shown in one case. The points —x—x— are those used to determine the position of the peak, as described in the text.

The position of the peak of each distribution was determined in the following way. The experimental points were plotted, and a smooth curve drawn through them. The mean abscissa values for pairs of points corresponding to a series of ordinate values throughout the curve were found, and a smooth curve drawn through these mean points. The locus thus formed intersected the top of the curve at a point whose abscissa was taken to be the position of the peak. In this way, the peak position was obtained from the distribution as a whole, and was not strongly dependent on the few points about the top of the peak. Thus the estimated error in locating any particular peak was generally less than $\frac{1}{2}$ per cent. and for the thinnest foils was of the order of 0.1 per cent. The displacement of the peak for each of the absorbers was found as the mean of the displacements in the three runs. Some typical experimental distributions are given in Figure 4.

IV. COMPARISON WITH THEORY

The distribution obtained from the incident beam (Fig. 4) has a width at half height of 9 per cent. This is made up of the incident energy spread, which was 2.3 per cent. plus the unknown width of the counter resolution curve. Assuming that the widths combine in the same way as standard deviations, this would give a width for the counter resolution curve of approximately 8.7 per cent.

Since the width of the theoretical energy loss curves was of the order of 10 per cent. it was necessary to include the effect of energy spread and counter resolution when comparing theory with experiment. Strictly, the energy spread should first be folded with the energy loss distribution and then the resulting curve folded with the counter resolution curve. However, since the energy spread was small compared with either the energy loss distribution or the counter resolution, it was theoretically satisfactory, and much more convenient, to fold the "energy spread plus counter resolution" curve with the theoretical energy loss distribution. Actually, a Gaussian approximation to the energy spread plus counter resolution curve was used in the folding process.

After the calibration described in Section III (b) had been carried out, the complete electron peaks at the various lower energies were investigated, so that information could be obtained about their widths. It was found that the widths in volts (i.e. *not* converted to an energy scale from the calibration curve) were constant, within the experimental error. This somewhat simplified the folding process. If the theoretical distribution of the energies of the particles emerging from the foil is $\varphi(E)dE$ and if particles of energy E give rise to a pulse height distribution $g(V-v)dV$, where the (E, v) relation is known from the calibration curve, the expected pulse height distribution is

$$f(V)dV = \int_0^\infty \varphi(E)g(V-v)dE = \int_0^\infty F(v)g(V-v)\frac{dE}{dv}dv dV,$$

where the (v, E) relation is expressed in the form $E=E(v)$, and $F(v)=\varphi(E(v))$. In this expression, $g(V-v)$ is taken to be a Gaussian of constant variance.

In this way, the Landau distribution was found for each absorber, and the Monte Carlo and Yang distributions for each of the three thickest foils. From these folded distributions values of ΔV , the difference between pulse height at the peak, and the pulse height for an energy of 1 MeV as found from the (E, v) calibration were obtained. The same quantity was obtained from the experimental curves. It should be noted that this is not equal to $\Delta V'$, the difference between the pulse height of the peak for each absorber, and the pulse height for the peak obtained with the blank hole in front of the crystal, because of the presence of the thin foil in the crystal mounting. $\Delta V'$ was the quantity obtained from each run (Section III (c)). Because the most probable energy loss is very nearly a linear function of thickness for small thicknesses, the voltage corresponding to the energy loss (3 keV) in the foil of the crystal mounting was obtained from the (E, v) curve, and added, in each case, to the mean value of $\Delta V'$, to give the experimental ΔV . The thickness of the foil of the crystal mounting was included in the thicknesses given for the absorbing foils as quoted in Table 1.

The results are shown in Table 1 and Figure 5. For the three thickest absorbers, the experimental and theoretical widths of the pulse height distributions are shown in Table 2. Figure 6 compares the experimental distribution for 0.156 g/cm^2 with the appropriate theoretical ones.

TABLE 1
EXPERIMENTAL AND THEORETICAL PEAK DISPLACEMENTS IN VOLTS FOR VARIOUS THICKNESSES
OF ABSORBER

The numbers in parentheses are the actual thicknesses used in the theoretical calculations

Thickness (mg/cm ²)	Experimental	Landau	Monte Carlo using		Landau-Yang
			Molière	Rossi- Greisen	
13.8	0.7 ± 0.1	(10) 0.4			
30.0	1.6 ± 0.1	(26) 1.3			
38.7	2.0 ± 0.2	(35) 1.5			
55.1	3.1 ± 0.2	(51) 2.3			
86.2	4.7 ± 0.2	(83) 3.7			
104.9	6.2 ± 0.3	(101) 4.8	(105) 5.7	6.3	(105) 5.7
156.3	9.2 ± 0.3	(152) 7.3	(158) 8.8	9.5	(162) 9.4
218.3	14.0 ± 0.5	(216) 11.0	(210) 12.4		(210) 13.5

The difference between the experimental distribution and the estimated background was integrated over the voltage range for the resolution curve and the three thicker foils to obtain an estimate of the fraction of the incident

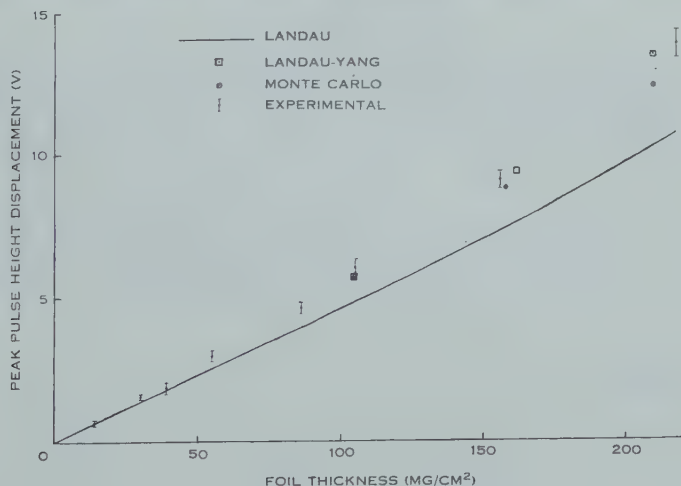


Fig. 5.—Experimental and theoretical peak displacements.

beam absorbed in the foils. Because it was difficult to estimate the tail, these calculations were not precise and all that can be concluded is that no absorption was observable in foils up to 0.158 g/cm^2 while for the 0.215 g/cm^2 foil absorption was of the order of 10 per cent. This is consistent with the Monte Carlo calcu-

lation which predicts percentage absorptions of 0.2, 1.5, and 8.4 respectively, but is not in accord with the values which can be interpolated from the absorption curves of Marshall and Ward (1937), viz. 4, 10, and 23 per cent.

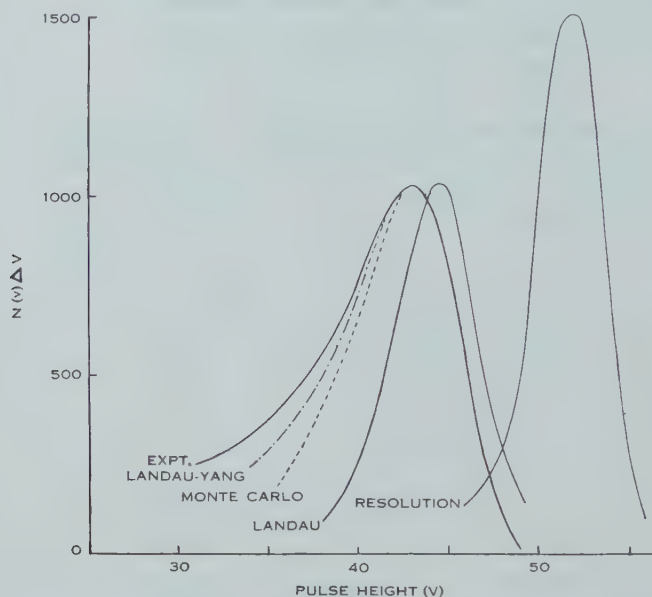


Fig. 6.—Experimental distribution compared with theoretical curves folded with resolution curve.

V. DISCUSSION

The agreement with the appropriate theory is generally satisfactory. For thicknesses less than about 50 mg/cm², the peak pulse height differences agree with the values predicted by the Landau theory. This has been demonstrated with more accuracy elsewhere (Chen and Warshaw 1951). For the region in

TABLE 2
WIDTHS, IN VOLTS, OF EXPERIMENTAL AND THEORETICAL PULSE HEIGHT DISTRIBUTIONS FOR THREE THICKNESSES OF ABSORBER
The numbers in parentheses are the actual thicknesses used in the theoretical calculations

Thickness (mg/cm ²)	Experimental	Monte Carlo	Landau-Yang
104.9	5.5 ± 0.3	(105) 5.6	(105) 5.6
156.3	6.7 ± 0.5	(158) 6.6	(162) 6.8
218.3	8.7 ± 1.0	(210) 9.2	(210) 10.1

which the effect of scattering becomes appreciable, the Landau theory underestimates the energy loss, satisfactory values of which can be obtained by combining the Landau theory with a theory of multiple scattering, even though the energy losses concerned may no longer be a small fraction of the incident energy.

As far as the method of combining the two theories is concerned, our results tend to favour the Landau-Yang folded distribution, rather than those obtained by the Monte Carlo (Molière) calculation. For the two foils to which it was applied, the Monte Carlo (Rossi-Greisen) distribution gave values of most probable ΔV which agreed very well with experiment. These values were not obtained by direct folding, however, but by increasing ΔV (Molière) in the ratio of the peaks of the unfolded curves.

The widths shown in Table 2 do not enable us to decide between the Monte Carlo and Landau-Yang distributions, and we would not expect them to do so, because they are, in effect, the widths of the counter resolution curve with some broadening due to the folding with energy loss distributions. However, from Figure 6 there is once more this tendency to favour the Landau-Yang distribution, as well as a clear broadening beyond that predicted by Landau theory alone.

Thus we find that the Landau-Yang distributions calculated by the techniques described in I, can be expected to be reasonably accurate. In particular, the simple expressions derived there for estimating the most probable energy loss and width of the distribution without going through the folding process, can be used with confidence.

VI. ACKNOWLEDGMENTS

We wish to thank Professor L. H. Martin for his advice and guidance throughout the course of this work; also Associate Professor C. B. O. Mohr, Mr. K. C. Hines, and Mr. D. F. Hebbard for useful discussions.

VII. REFERENCES

- BLUNCK, O., and LEISEGANG, S. (1950).—*Z. Phys.* **128**: 500.
BORKOWSKI, C. J., and CLARK, R. L. (1953).—*Rev. Sci. Instrum.* **24**: 1046.
CHEN, J. J. L., and WARSHAW, S. D. (1951).—*Phys. Rev.* **84**: 355.
DEUTSCH, M., ELLIOT, L. G., and EVANS, R. D. (1944).—*Rev. Sci. Instrum.* **15**: 178.
GOUDSMIT, S. A., and SAUNDERSON, J. L. (1940).—*Phys. Rev.* **57**: 24.
HEBBARD, D. F., and WILSON, P. R. (1955).—*Aust. J. Phys.* **8**: 90.
LANDAU, L. (1944).—*J. Phys., Moscow* **8**: 201.
MARSHALL, J. S., and WARD, A. G. (1937).—*Canad. J. Res. A* **15**: 39.
YANG, C. N. (1951).—*Phys. Rev.* **84**: 599.

MICROWAVE AND METRE WAVE RADIATION FROM THE POSITIVE COLUMN OF A GAS DISCHARGE

By L. W. DAVIES* and E. COWCHER†

[Manuscript received September 16, 1954]

Summary

Experiments are described in which the intensity of radio-frequency "noise" radiation from the positive column of low pressure discharges in neon was measured at 3000 Mc/s and at 200 Mc/s. Measurements were made by inserting the positive column in a resonant cavity and comparing the cavity output with that from a noise generator of known intensity; corrections were made for contributions to the noise radiation from dissipative elements in the cavity at room temperature. The mean energy of the electrons in the discharge was calculated from measured values of the electric field.

As the discharge current was lowered the noise temperature T' at 3000 Mc/s increased above the calculated value of the electron temperature for an assumed Maxwellian distribution of electron velocities; this phenomenon indicated a change in the distribution function as the electron density was reduced. The thermal level of noise radiation from a plasma with Druyvesteyn distribution of electron velocities was calculated and compared with results for low values of the discharge current. For high values of the current there was good agreement between the values of T' at 3000 Mc/s and the electron temperature only when the pressure was sufficiently high; for lower pressures T' was less than the electron temperature.

In a limited range of discharge conditions the measured values of T' at 200 Mc/s were considerably greater than the thermal level. Experiments were carried out to ascertain the source of this enhanced radiation.

I. INTRODUCTION

An ionized medium in equilibrium at temperature T , and of sufficient optical depth, will emit radiation of all frequencies with an intensity equal to that of a black body at the same temperature. In other words, at radio frequencies, the available noise power from an ionized medium in equilibrium at temperature T amounts to kT per unit bandwidth. In general, the positive column of a gas discharge is not in equilibrium, and we specify the radio-frequency radiation from a discharge by means of an equivalent noise temperature T' ; it is such that the available noise power per unit bandwidth is kT' , where T' may be a function of frequency.

Microwave "noise" radiation from the positive column was first observed by Mumford (1949); it has been investigated in more detail by several writers (Easley and Mumford 1951; Johnson and DeRemer 1951; Knol 1951). In general, the equivalent noise temperature T' has been found to be comparable

* Division of Radiophysics, C.S.I.R.O., University Grounds, Sydney.

† Division of Electrotechnology, C.S.I.R.O., University Grounds, Sydney.

with the temperature T_e of the electrons in the discharge; however, it was thought worth while to carry out a more extensive investigation than had been made previously, particularly at low values of the discharge current. Relationships between electron temperature, equivalent noise temperature, and collision frequency of the electrons in the discharge are important in connexion with the development of discharges as standard noise signal generators; it is also of interest to study such relationships in the light of theories of the origin and transfer of radio-frequency radiation in an ionized medium (Smerd and Westfold 1949) with their particular application to radio-astronomical observations.

This paper describes experiments in which the noise radiation from the positive column of discharges in neon was observed at frequencies of 3000 and 200 Mc/s, over a wide range of discharge conditions.

Preliminary results were reported at the 10th General Assembly of the International Union of Scientific Radio (U.R.S.I.), Sydney, 1952.

II. CONTRIBUTION OF THE POSITIVE COLUMN TO THE OBSERVED RADIATION

When the positive column is inserted in a resonant cavity, it may be considered as a load at a temperature equal to the equivalent noise temperature of the discharge. To determine the contribution of this load to the observed noise from the cavity, we consider the equivalent circuit of the cavity and impedance transformer as set down in Figure 1 (a), in which reactances introduced

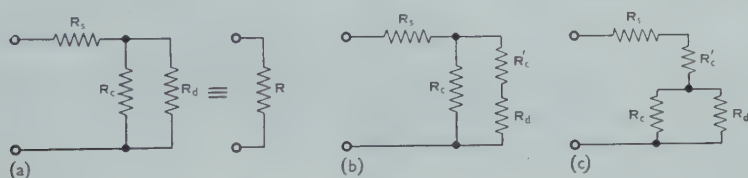


Fig. 1.—Equivalent circuits of resonant cavity containing discharge tube.

by the discharge have been neglected (see Section IV); the resistances R_s , R_c , and R_d correspond respectively to losses in the transformer, cavity, and discharge. Regarded as generators of noise, the resistances R_s and R_c are at room temperature T_r , and the resistance R_d at the effective noise temperature of the discharge, T' . We denote by T_n the observed noise temperature of the cavity when matched to the transmission line.

From a consideration of the relative losses in the resistances of the circuit given in Figure 1 (a), it is easily shown that, where R is the resistance of the circuit of Figure 1 (a),

$$T_n = T' \left[\frac{1}{R} \cdot \frac{R_c^2 R_d}{(R_c + R_d)^2} \right] + T_r \left[\frac{1}{R} \left(R_s + \frac{R_c R_d^2}{(R_c + R_d)^2} \right) \right]. \quad \dots (1)$$

In order to determine T' from observed values of the available noise power from the cavity, standing-wave measurements are made of the impedance presented by the cavity with discharge off, i.e. with resistance R_d removed from the equivalent circuit; these are carried out at the resonant frequency of the

cavity (voltage standing-wave ratio, V.S.W.R. = σ_r) and at a frequency well removed from resonance (V.S.W.R. = σ_0). Then we have (Slater 1950) :

$$\left. \begin{aligned} 1 &= r_s + (r_d^{-1} + r_c^{-1})^{-1}, \\ \sigma_r &= r_s + r_c \geq 1, \\ \sigma_0 &= r_s^{-1}, \end{aligned} \right\} \dots\dots\dots (2)$$

where $r_s = R_s/R$, $r_c = R_c/R$, $r_d = R_d/R$, and we assume the cavity to be matched to unit standing wave ratio with discharge running. From (1) and (2) it follows that

$$T' = \frac{\sigma_r \sigma_0 - 1}{(\sigma_r - 1)(\sigma_0 - 1)} T_n - \frac{\sigma_r + \sigma_0 - 2}{(\sigma_r - 1)(\sigma_0 - 1)} T_r.$$

In general, the second term may be neglected, since $T_n \gg T_r$ and $\sigma_0 \gg 1$, giving

$$T' = \{(\sigma_r - \sigma_0^{-1})(\sigma_r - 1)^{-1}(1 - \sigma_0^{-1})^{-1}\} T_n. \dots\dots\dots (3)$$

If the resistive load corresponding to the discharge losses is not completely in shunt with the cavity losses R_c , the equivalent circuit of the impedance presented by the cavity and matching transformer may take the forms shown in Figures 1 (b) and 1 (c), where R'_c represents losses in the cavity at room temperature which are not shunted by R_d . In either of these cases it may be shown that

$$T' \geq \{(\sigma_r - \sigma_0^{-1})(\sigma_r - 1)^{-1}(1 - \sigma_0^{-1})^{-1}\} T_n. \dots\dots\dots (4)$$

Such a case may arise, for example, when losses occur in the glass walls of the cylindrical discharge tube and the axis of the tube is not parallel to the electric field vector in the cavity (cf. Section III (a)).

III. EXPERIMENTAL TECHNIQUE

(a) Measurement of T' at 3000 Mc/s

A block diagram of the circuit used is shown in Figure 2. The resonant cavity consisted of a section of waveguide, short-circuited at one end and provided with two sliding stubs of variable depth at the other, with the discharge tube mounted in the E -plane of the TE_{01} mode (H_{01}) and inclined at an angle of 15° to the wider face of the waveguide. Lengths of cylindrical copper tubing, of diameter 2.5 cm, surrounded that portion of the tube extending outside the guide; they acted as a waveguide beyond cut-off for the 10-cm radiation, and were calculated to provide approximately 100 db attenuation for radiation entering the guide from the electrode regions of the discharge tube.

The noise temperature of the cavity, T_n , was determined by comparing its output with that of an argon gas discharge tube ($T = 10,250^\circ \text{K}$), which in turn had been calibrated by direct reference to a hot, dissipative waveguide load (cf. Section III (b)). The procedure was to connect the "cavity" section of guide to the impedance measuring equipment and match the cavity to the waveguide (V.S.W.R. ≤ 1.03) by means of the sliding-stub transformer. The noise-measuring receiver and image-frequency rejection cavity were accurately

aligned with the signal generator, and thus with the frequency at which the cavity was matched, and the cavity section of the waveguide transferred to the noise-measurement gear. The image rejection cavity was necessary because the different bandwidths of the two noise sources gave rise to different degrees of mismatch, and hence of noise energy transfer, at the image frequency. The transmission of the cavity at the image frequency was 42 db down on that at the signal frequency. For most measurements the bandwidth of the intermediate-frequency (30 Mc/s) amplifier of the noise-measuring receiver was 500 kc/s; for the lowest pressure investigated, and for low values of the discharge current, however, the bandwidth of match was found to be so narrow that a second

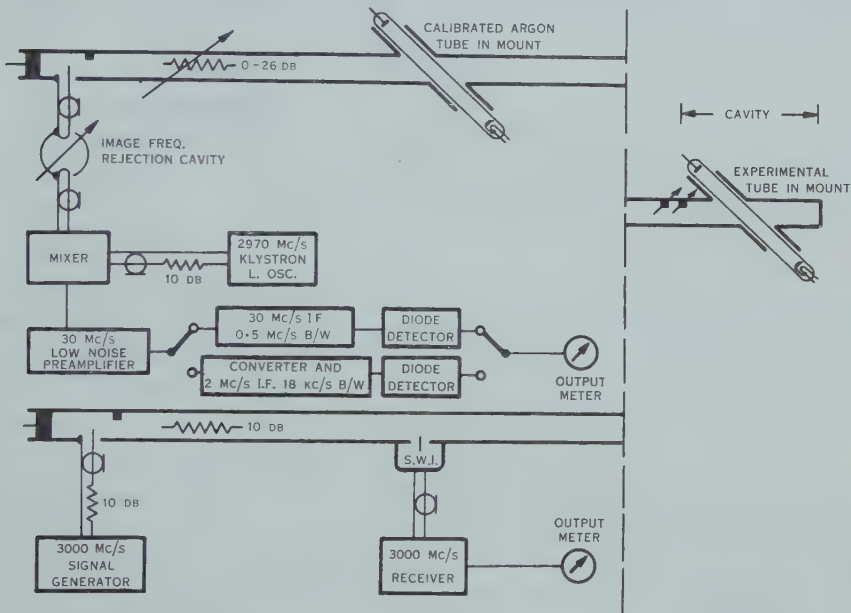


Fig. 2.—Experimental arrangement for measurement of noise temperature of discharge at 3000 Mc/s.

conversion to 2 Mc/s was employed, with a bandwidth of 18 kc/s between points 3 db down from the central frequency. Frequency stability was realized by using regulated voltage supplies throughout the electronic equipment. The noise figure of the receiver, including image rejection cavity, was of the order of 12 db.

The value of T_n was determined by measuring firstly M_n , the ratio of receiver output powers with cavity and room-temperature attenuator separately contributing noise signal to the input, followed by a measurement of M_s , the corresponding ratio for the substandard noise source. It can be shown that

$$T_n = \frac{M_n - (1 - F^{-1})}{M_s - (1 - F^{-1})} \cdot T_s, \quad \dots \dots \dots (5)$$

where F is the noise figure of the receiver, obtained from the formula

$$F = \frac{1}{(M_s - 1)} \left\{ \frac{T_s}{T_r} - 1 \right\}, \quad \dots \dots \dots (6)$$

and T_s is the noise temperature of the substandard noise source. Several determinations were made of M_n and M_s for each set of discharge conditions investigated.

With discharge off, the cavity was then transferred to the impedance measurement equipment where the relevant standing-wave ratios σ_r , σ_0 , were determined. Since the presence of the discharge generally gave rise to a small change in resonant frequency of the cavity with discharge on and off, σ_r was determined by measuring V.S.W.R. as a function of frequency (cf. Fig. 3) and

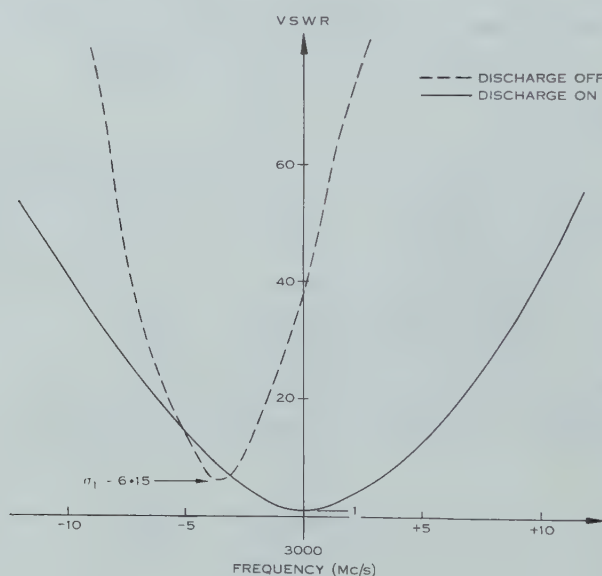


Fig. 3.—Voltage standing-wave ratio of cavity as a function of frequency near 3000 Mc/s. Experimental conditions: $p_0 = 2.80$ mm Hg, discharge current 10 mA.

taking σ_r to be the V.S.W.R. at resonance, with discharge off. At 3000 Mc/s σ_0 was found to be greater than 200, and accordingly σ_0^{-1} was neglected in equation (3), giving

$$T' = \{ \sigma_r / (\sigma_r - 1) \} T_n. \quad \dots \dots \dots (7)$$

By determination of the position of the standing-wave minimum as a function of frequency at the same time, it was verified that the Q of the unloaded cavity (i.e. with discharge off) was less than the external Q (see Slater 1950), so that the equivalent circuits of Figure 1 and the derivation of equation (3) are valid.

(b) Calibration of 3000 Mc/s Substandard Noise Source

The determination of the noise temperature of the substandard argon discharge tube was performed by comparing its noise output with the noise

output from a wedge-shaped waveguide termination heated uniformly to a temperature of approximately 450 °K. A second termination was used to obtain a room temperature reference.

The large difference between the temperature of the thermal standard (450 °K) and the effective temperature (10,000 °K) of the discharge, led to considerable experimental difficulties: the most troublesome of these were:

1. The receiver was particularly sensitive to mismatch of the room temperature termination and the 450 °K thermal standard. It may be shown (F. F. Gardner, personal communication) that, when the equivalent input temperature of the receiver is greater than the temperature of the input termination, a change of 1 per cent. in the resistive component of the termination can give rise to a relatively greater apparent change (estimated to be 4 per cent. in this case) in the equivalent temperature of the termination.
2. The difference in noise power from the two sources required that either the variation of the gain of the receiver with power output be known, or that the two sources be brought to the same intensity by means of an attenuator.

In order to overcome the first of these difficulties a telescopic waveguide was placed at the input of the receiver. The impedance of the terminations was adjusted until a negligible change in output occurred as the length of the input guide was changed.

Several different arrangements of noise measuring equipment were used to overcome the second of these difficulties.

Initially a straight superheterodyne receiver capable of receiving both the image and signal frequencies was used. The gain versus power output law of the I.F. amplifier and second detector was determined by means of a 30 Mc/s noise diode. The crystal mixer was assumed linear for the small signals involved. Owing to the reception at the image frequency difficulty was experienced in obtaining a suitable match with the various terminations.

A more satisfactory experimental set-up (Fig. 4) used a receiver with a Dicke modulating system (Dicke 1946). The increased stability of the output signal level which was obtained with this arrangement permitted the use of an image rejection cavity in front of the mixer. This eliminated the matching difficulties experienced with the straight receiver. However, since this type of receiver is essentially non-linear, a calibrated waveguide attenuator was placed between the discharge tube mount and the input of the receiver. By this means the equivalent noise temperature was reduced to the same level as that of the thermal standard.

As a check on the above measurement, the noise power output versus plate current law of a coaxial noise diode was measured at 3000 Mc/s, using the previously calibrated "straight" receiver. This was possible since the only departure from linearity of the coaxial diode occurred at very high noise outputs.

The noise diode was then used as a transfer instrument, being calibrated against the hot load at low intensities and compared with the discharge at high intensities.

The results of each of these measurements are given in Table 1 together with the estimated errors.

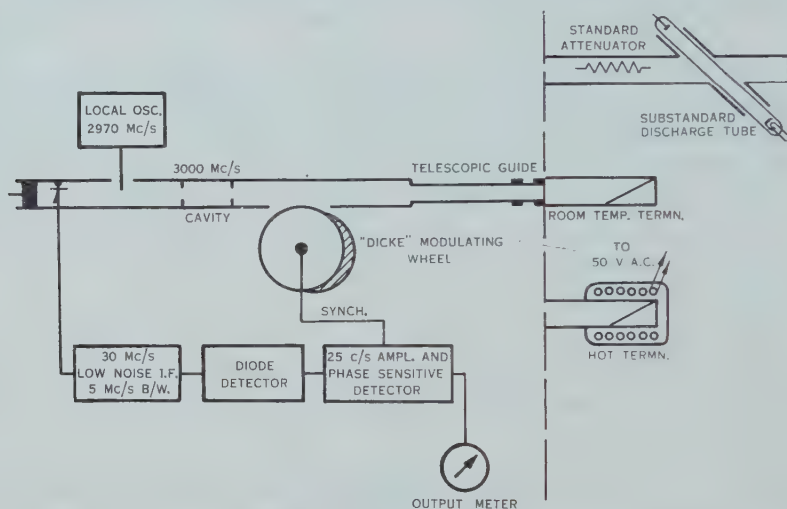


Fig. 4.—Measurement of noise temperature at 3000 Mc/s of substandard discharge tube.

(c) Measurement of T' at 200 Mc/s

Measurements of the noise radiated by the positive column at 200 Mc/s were made by inserting the positive column into the high-field region of a re-entrant, tunable resonant cavity, as shown in the block diagram of Figure 5. In this case also, the portion of the positive column projecting beyond the resonant cavity was surrounded by cylindrical guide beyond cut-off, to prevent radiation generated in the electrode regions passing into the cavity.

TABLE 1
DETERMINATION OF SUBSTANDARD NOISE TEMPERATURE AT 3000 mc/s

Method of Calibration	Equivalent Noise Temperature ($^{\circ}\text{K}$)	Estimated Error (%)
1. "Straight" receiver	10,300	± 10
2. Dicke system and attenuator	10,200	± 5
3. Dicke system and noise diode	10,300	± 5

The effective noise temperature of the cavity was determined by comparison with a diode noise generator whose internal reactance was tuned out at 200 Mc/s (Love 1948) and impedance matched to the characteristic impedance (75Ω) of the transmission cable. With discharge on, the cavity was attached to the impedance measuring gear and matched to unit standing-wave ratio

(V.S.W.R. ≤ 1.06) by means of the coarse and fine capacitive tuning of the cavity and the degree of coupling of the loop. The signal frequency used for matching was accurately aligned beforehand with that of the noise measuring receiver, to which the coaxial lead from the cavity was transferred after matching. The noise measuring receiver was made up of a tuned 200 Mc/s preamplifier, followed by intermediate-frequency amplification at 30 Mc/s and at 2 Mc/s, with an effective bandwidth of 18 kc/s. The noise figure of the receiver was of the order of 10 db.

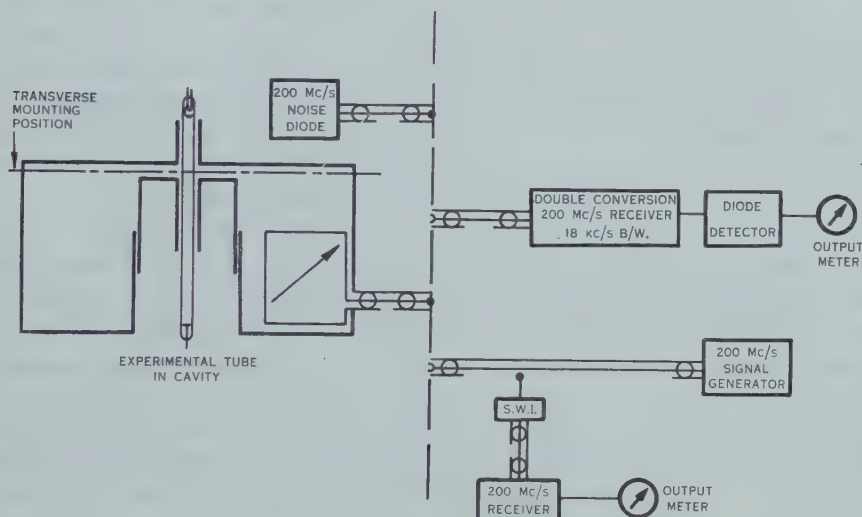


Fig. 5.—Measurement of noise temperature of discharge at 200 Mc/s.

As in the measurements at 3000 Mc/s, the noise temperature T_n of the cavity at 200 Mc/s was determined by observing the receiver output power ratios, M_n and M_s , and inserting these values in equations (5) and (6). In this case the noise temperature T_s of the standard noise source (diode generator) is given by

$$T_s = eIR/2k + T_r \quad \dots \dots \dots (8)$$

For the experimental conditions $I = 20 \text{ mA}$, $R = 75 \Omega$, and $T_r = 300^\circ \text{K}$; this gives $T_s = 9010^\circ \text{K}$.

With discharge off, standing-wave measurements were then made at the resonant frequency of the cavity and at a frequency well removed from resonance, and the noise temperature T' of the discharge calculated from equation (3). Owing to losses in the coaxial cable connecting cavity and receiver input, σ_0 was found to be of the order of 20 db and was not negligible at this frequency.

(d) Determination of Discharge Parameters

The measurements at 3000 and 200 Mc/s were made on the positive column of discharges in neon. Several tubes were fabricated of Pyrex glass, filled at pressures in the range 0.5–10 mm Hg and sealed off. The positive column in each case was 1.0 cm in diameter and of approximate length 80 cm; each tube

was provided with a thermionic cathode, which was mounted in a buffer volume of neon at the sealing-off pressure. Care was taken to remove foreign gases thoroughly before sealing-off, and no significant changes were detected in the electrical characteristics of the tube during its experimental life. Spectroscopic examination of the radiation from the discharge was also used as a check on the purity of the gas.

A determination of the mean energy of the electrons in the discharge, available over the whole range of discharge conditions, is made possible by the existence of theoretical relationships between the mean energy of the electrons and E/p_0 , the electric field strength divided by the pressure of neon at 0°C . If there is sufficient interaction between the electrons in the discharge, they will have a Maxwellian distribution of velocities with electron temperature T_{e1} , and Mierdel (1938) has given the relationship between T_{e1} and E/p_0 in this case. For the case in which there is negligible interaction between the electrons, Druyvesteyn has derived a theoretical distribution of electron velocities (given in equation (A10) of Appendix I), and the relationship between T_{e2}^* and E/p_0 for this distribution in neon has been given by Druyvesteyn and Penning (1940). These writers also show that there is good agreement between experimental values of T_e and those predicted theoretically, for both sets of conditions.

The value of the electric field strength for insertion in these theoretical relationships was taken to be the longitudinal electric field in the positive column. Because of the great length of the positive column, this was determined with sufficient accuracy by dividing the potential drop across the discharge by its total length; it was assumed that the cathode fall at the thermionic cathode was negligible. Measurement of the voltage gradient in one case by the use of two axial probes showed this to be a valid approximation. These measurements enabled estimates to be made of the electron temperature T_e for the two separate cases in which either a Maxwellian or a Druyvesteyn distribution was assumed to exist in the plasma.

IV. EXPERIMENTAL RESULTS

Measured values of the effective noise temperature at 3000 Mc/s and at 200 Mc/s are shown in Figures 6–8 as a function of the current in each of four discharge tubes.

For low values of the current in the two tubes of lowest pressure there were, initially, oscillations in the discharge current at approximately 2 kc/s. These oscillations gave rise to anomalous standing-wave patterns at 3000 Mc/s, which were superpositions of the standing-wave patterns corresponding to the varying load presented by the discharge. It was found possible to suppress these oscillations in the case of one tube by inserting a resistor in the tube circuit at the anode, and no difficulty was then experienced in matching the discharge to V.S.W.R. ≤ 1.03 . However, the tube of lowest filling pressure (0.495 mm Hg at 0°C) could not be prevented from oscillating below currents of 150 mA.

* In the non-Maxwellian case, the electron temperature is defined (Chapman and Cowling 1951) by: mean energy of electrons $= 3kT_e/2$.

Corrections to the observed noise level of the cavity for cavity losses were given by equation (7). At 3000 Mc/s the correction factor $\sigma_r/(\sigma_r-1)$ ranged from unity for currents greater than approximately 20 mA to values as high as 5 as the discharge current was reduced to a few hundred microamperes. Thus

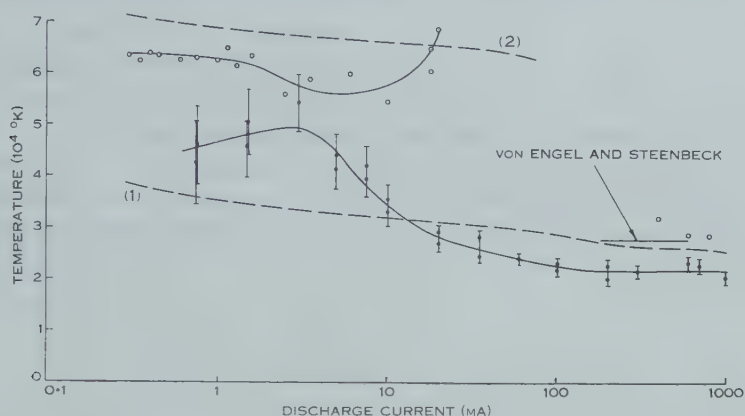


Fig. 6.—Measurements of effective noise temperature of discharge for $p_0=2.80$ mm Hg: \cdot 3000 Mc/s; \odot 200 Mc/s; \circ probe measurements of electron temperature.

the estimated experimental errors are greater at lower values of the current, as shown in Figures 6–8; these errors do not include possible errors in the determination of the temperature of the substandard, which amounted to ± 5 per cent.

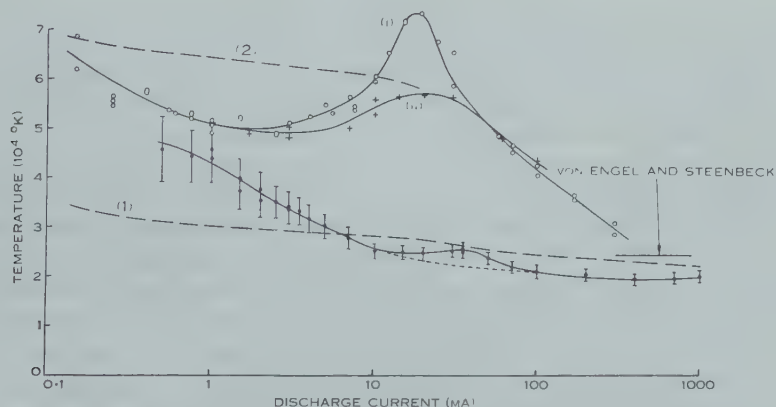


Fig. 7.—Measurements of effective noise temperature of discharge for $p_0=4.75$ mm Hg: \cdot 3000 Mc/s; \odot 200 Mc/s, with axis of discharge tube parallel to E -vector in cavity; $+$ 200 Mc/s, with discharge tube in transverse mounting position.

The observed noise levels at 200 Mc/s were corrected for cavity and series losses by use of equation (3). The correction factor ranged approximately from 1.20 at higher values of the current investigated, up to 2.5 for low currents. The estimated errors in the measurements at 200 Mc/s amounted to ± 10 per cent.

In both discharge tubes investigated at 200 Mc/s it was observed that the value of T' at first increased rapidly as the discharge current rose above a value of approximately 10 mA. In the case of the tube with filling pressure 4.75 mm Hg (Fig. 7), the measurements were extended to higher values of the discharge current; it was found that the value of T' passed through a maximum at about 20 mA, and decreased monotonically for increase in current above this value. Also shown in Figure 7 are some results obtained with the discharge tube mounted transversely in the resonant cavity, in the position shown in Figure 5.

Further experiments were carried out with the tube of Figure 7 in order to investigate the considerable enhancement of the values of T' at 200 Mc/s over those at 3000 Mc/s, for the range of discharge current 1–100 mA. When a signal at 200 Mc/s was fed into the cavity and a pick-up loop placed in the vicinity of the positive column outside the cavity, the guides beyond cut-off

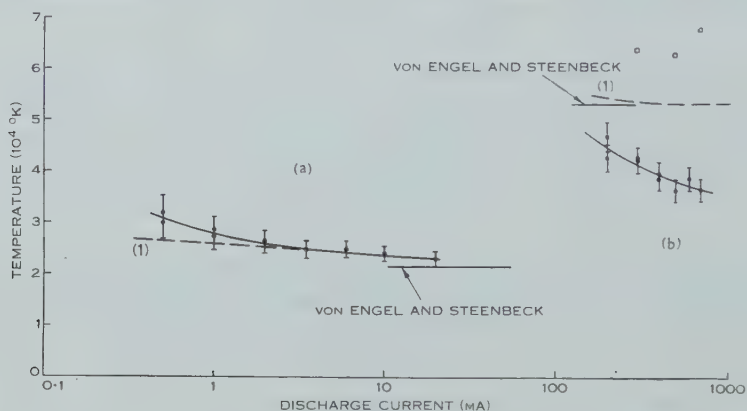


Fig. 8.—Measurements of effective noise temperature at 3000 Mc/s of discharge; (a) $p_0=9.42$ mm Hg, (b) $p_0=0.46$ mm Hg. \circ Probe measurements of electron temperature.

surrounding the discharge as it passed into the cavity were found to conduct 200 Mc/s radiation as the current rose above 5 mA. This conduction presumably took place with the discharge acting as the centre conductor of a coaxial line (Rosen 1949). Measurements were made of the attenuation of the signal along the positive column when used in conjunction with an earthed, plane conductor—as a strip transmission line; it amounted to 2 db/cm for a current of 10 mA in the discharge tube of Figure 7, and decreased with increase in current. Thus the attenuation provided by each copper tube surrounding the positive column as it entered the cavity was estimated to decrease from approximately 100 db to less than 15 db as the discharge current increased through the critical value for coaxial transmission.

With the discharge current at 20 mA, the noise level at 200 Mc/s could be increased by amounts up to 10 per cent. by bringing up a magnet to the cathode region of the discharge, or by decreasing the filament supply voltage. By moving the tube longitudinally with respect to the cavity, so that the distance of the cathode from the cavity was reduced by approximately 60 cm, the noise level

from the cavity could be increased monotonically up to a maximum of 1.14 times the level observed with maximum separation; this experiment was carried out with the discharge current at 20 mA. Measurements were made of the 200 Mc/s noise generated (by fluctuations in the discharge current) in a $75\ \Omega$ resistor placed in series with the discharge tube; in the range of discharge current 20–100 mA the effective noise temperature of the resistor increased from 3000 to 6000 °K.

From measured values of the longitudinal electric field in the discharge, values of the electron temperature were calculated as discussed in Section III (*d*). The calculations were carried out for two assumed distribution functions, namely, Maxwell's and Druyvesteyn's. Calculations given in Appendix I show that the effective noise temperature of a plasma with a Druyvesteyn distribution, of mean energy $\bar{\epsilon} = 3kT_{e2}/2$, is $0.874 T_{e2}$ when the radio-frequency radiation arises as bremsstrahlung in collisions between electrons and positive ions; the values of T_{e1} for a Maxwellian distribution and of $0.874 T_{e2}$ for a Druyvesteyn distribution are given in Figures 6–8 as a function of the discharge current.

Probe measurements of electron temperature were attempted, the results being shown in Figures 6 and 8. The measurements were not felt to be particularly reliable, because of the small diameter of the positive column and the resultant disturbance to the properties of the discharge in its immediate vicinity by the probe. For comparison, theoretical values of the electron temperature, calculated from an expression given by von Engel and Steenbeck (1934), are also shown in Figures 6–8.

V. DISCUSSION

As an introduction to the discussion of the results presented in Section IV, we consider the expected level of thermal radio-frequency radiation from the plasma of a discharge in neon. For a sufficiently high value of the electron density, the electrons have a Maxwellian distribution of velocities (with electron temperature T_{e1}) due to the interaction between electrons, while for values of the electron density so low that electron interaction is negligible, we assume the distribution to have the form derived by Druyvesteyn, with mean energy equal to $3kT_{e2}/2$. In Appendix I it is shown that the values of the equivalent noise temperature T' of the plasma in the two cases are respectively T_{e1} and $0.874 T_{e2}$, when the radiation is emitted as bremsstrahlung in the collisions between electrons and positive ions. If measurements are made of the electric field strength in the positive column and values of the mean energy of the electrons calculated on the basis of one or the other distribution as the true one, we arrive at two curves describing the relationship between the value of T' and the current in the discharge, to which the electron density is related. These curves are shown schematically in Figure 9; curves (1) and (2) give the values of T_{e1} and $0.874 T_{e2}$ respectively. If the electrons have a Maxwellian distribution of velocities for values of the discharge current greater than I_1 , and electron-electron interaction is negligible for values less than I_2 , the value of T' is equal to T_{e1} for values of current greater than I_1 , and $0.874 T_{e2}$ for values less than I_2 . Between these values of the discharge current the electron velocity distribution

undergoes a transition from one form to the other, and the equivalent noise temperature at all radio frequencies is accordingly expected to have the form shown in Figure 9 over the full range of discharge current.

The results presented in Figures 6–8 are in qualitative agreement with the theory outlined above, and we shall take the point of view that the observed increase (Figs. 6, 7) in T' at 3000 Mc/s to values greater than T_{e1} indicates a departure from a Maxwellian distribution of the electron velocities as the discharge current, and with it the electron density, decreased from the highest values. There are several features of quantitative departure from the theoretical picture, particularly in the measurements at 200 Mc/s, and we shall discuss them in turn.

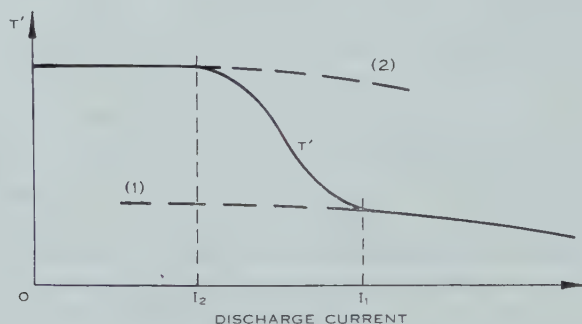


Fig. 9.—Theoretical value of the noise temperature of the positive column, T' , as a function of current in the discharge.

At values of the discharge current less than 3 mA in Figure 6 and 1.5 mA in Figure 7, the correction factor for cavity losses at 3000 Mc/s was in excess of 2. Consequently the values of T' at 3000 Mc/s shown in Figures 6 and 7 for these respective ranges of discharge current may be lower than the true values, in accordance with equation (4). This is the probable reason for the observed difference in the values of T' at 3000 and 200 Mc/s for the lowest values of the discharge current.

It is probable that the electron distribution function, in the absence of electron interaction, is not exactly that derived by Druyvesteyn; it is also possible that the effective noise temperature of a Druyvesteyn plasma may differ from $0.874 T_{e2}$ for other processes of emission and absorption of radiation than those considered in Appendix I. In the range of discharge conditions in which the electron density is so low that electron interaction is negligible the agreement between observed values of T' and the calculated value $0.874 T_{e2}$ may thus be poor.

A feature of the results shown in Figures 6–8 is the relationship between the noise temperature at 3000 Mc/s and the electron temperature in the region of higher discharge currents, where the electrons may be assumed to have a Maxwellian distribution of velocities. The values of T' and T_{e1} at different pressures are shown in Table 2, together with estimated values of the mean frequency with which an electron collides with atoms of the gas under the given conditions.

It is seen that the value of T' decreased progressively below T_{e1} as the discharge pressure was reduced, and that the collision frequency of the electrons simultaneously fell to values less than the frequency of observation ($3 \times 10^9 \text{ sec}^{-1}$). Johnson and DeRemer (1951) and Knol (1951) have also reported values of $T' < T_{e1}$, but it is not clear from their papers whether they have taken into account room temperature contributions to the microwave radiation from cavity losses.

If the radiation arises from discrete microscopic processes in the plasma, such as the bremsstrahlung emitted by electrons in their encounters with positive ions, there is no dependence of the effective noise temperature on the time rate at which these processes occur in a medium of infinite optical depth, other than that they should occur at random. The effective noise temperature for the radiation arising from microscopic processes involving the electrons in such a

TABLE 2

ELECTRON TEMPERATURE, NOISE TEMPERATURE AT 3000 Mc/s, AND ESTIMATED COLLISION FREQUENCY FOR MAXWELLIAN DISTRIBUTION OF ELECTRON ENERGIES

p_0 (mm Hg)	Current (mA)	T' (°K)	T_e (°K)	Z (sec^{-1})
9.42	10	23,800	23,800	$\sim 1.0 \times 10^{10}$
4.75	100	21,300	24,400	5.1×10^9
2.80	150	22,200	28,000	3.5×10^9
0.46	~ 700	36,400	53,000	7.4×10^8

medium will be the electron temperature T_{e1} , provided only that the motion of the other partners in the process (e.g. ions, atoms) may be neglected relative to that of the electrons. This is a valid assumption for almost all the electrons of a Maxwellian distribution in the positive column. The observed decrease in the value of T' at 3000 Mc/s below T_{e1} leads to the conclusion that there are mechanisms for the generation of radio-frequency noise in the positive column other than microscopic processes involving the electrons, and that these additional mechanisms have an effective noise temperature at 3000 Mc/s which may fall below T_{e1} . An example of such mechanisms might be the fluctuations in the current through the discharge which arise from the collisions of the charge carriers (electrons) with the atoms of the gas.

For the case in which the electrons have a Maxwellian distribution the results given here are not sufficiently comprehensive to compare the values of T' at 3000 and 200 Mc/s and thus to establish the significance of the decrease in collision frequency below $3 \times 10^9 \text{ sec}^{-1}$ as the value of T' at 3000 Mc/s decreased below T_{e1} (Table 2).

We now return to a consideration of the limiting values of discharge current below which the electrons in each tube first showed departures from a Maxwellian distribution of velocities. These values were derived from the measurements

of T' at 3000 Mc/s, shown in Figures 6-8; they were taken to be the values at which T' first showed a significant upward tendency in relation to the value of T_{e1} as the discharge current was reduced from the highest values. The limiting values of discharge current are given in Table 3, together with the corresponding estimated values of the axial electron density. The latter were calculated using values for the mobility of electrons in neon determined by Nielson (1936); it was assumed that the mean density of electrons in the positive column was equal to 0.7 times the axial density. Also shown in Table 3 are the relevant values of the parameter E/p_0 , which is a measure of the energy gained from the electric field by an electron between collisions. The results show that a Maxwellian distribution is established for smaller values of the electron density as the energy gained by an electron during one free path is decreased.

TABLE 3
ELECTRON DENSITY BELOW WHICH ELECTRON DISTRIBUTION FUNCTIONS SHOW
DEPARTURES FROM THE MAXWELLIAN

p_0 (mm Hg)	Limiting Current (mA)	Calculated Axial Electron Density (cm^{-3})	E/p_0 ($\text{V cm}^{-1} (\text{mm Hg})^{-1}$)
9.42	4	4×10^{10}	~ 0.7
4.75	10	7×10^{10}	1.37
2.80	30	1.4×10^{11}	1.93
0.46	≥ 700	$\geq 8 \times 10^{11}$	~ 9.0

Calculations on the interaction of electrons in a plasma (Cahn 1949) show that it should be sufficient to maintain a Maxwellian distribution for electron densities greater than 10^{11} – 10^{12} cm^{-3} , for particular values $E=1 \text{ V/cm}$ and electron mean free path $\lambda=10^{-2} \text{ cm}$. This corresponds in the case of neon to $E/p_0 \approx 0.2 \text{ V cm}^{-1} (\text{mm Hg})^{-1}$. The lowest value of E/p_0 investigated here was 0.7, for which the critical density was estimated to be $4 \times 10^{10} \text{ cm}^{-3}$. In view of the approximate nature of this value, it cannot be said that there is wide disagreement between these results and Cahn's theory. The observation that the critical density increased as the energy gained by an electron during one free path became larger (Table 3) is in accord with the idea that electrostatic interaction between electrons is primarily responsible for the establishment of a Maxwellian distribution.

We discuss finally the measurements of the effective noise temperature at 200 Mc/s, the results of which are shown in Figures 6 and 7. For discharge currents below 1 mA the value of T' is believed to be accounted for by the absence of electron interaction at low electron densities, so that the distribution of electron velocities is non-Maxwellian. For values of the discharge current above 1 mA we shall confine our discussion to the more extensive results of Figure 7.

The fact that a decrease in the attenuation for externally generated signals entering the cavity occurred at a discharge current of approximately 5 mA suggests that the increase in T' at 200 Mc/s, observed in the range 1–100 mA, may have been due in part to the penetration into the cavity of radiation generated at the electrodes. Labrum and Bigg (1952) have shown that the cathode of a discharge can act as a generator of radio-frequency noise at a high level. The observation of an increase in T' as the cathode moved closer to the cavity in this case is evidence that 200 Mc/s radiation from the cathode region probably played some part in the enhancement of the values of T' .

From measurements of the noise in the resistor placed in series with the discharge tube it was shown that there were fluctuations in the current in the discharge, with a component at 200 Mc/s which increased monotonically in the range 20–100 mA. This modulation may have originated in the electrode regions outside the resonant cavity. If it is assumed that the fluctuations in the discharge current at 200 Mc/s interact with the electric field in the cavity for the longitudinal position of the tube, the fluctuations will make a contribution to the noise level in the cavity. Experiments were carried out with the same discharge tube inserted in the cavity in such a position that this interaction was reduced, i.e. with the direction of current flow in the discharge normal to the electric field vector in the high-field region of the cavity: the noise level was in fact found to be reduced (Fig. 7, curve (ii)) by an amount comparable with the current noise developed in the series resistance. It is concluded that fluctuations in the discharge current at 200 Mc/s also contributed to the enhancement of T' in the range 1–100 mA.

There remains the possibility that the increase in values of T' at 200 Mc/s above those at 3000 Mc/s, in the range of discharge current 1–100 mA in Figure 7, was in part due to an intrinsic contribution to the noise level from the positive column; i.e. the plasma within the cavity generated noise in excess of the thermal level. This possibility is supported by a re-examination of the results at 3000 Mc/s, which show some evidence of a small maximum in the value of T' at a discharge current of 35 mA. This increase in the value of T' at 3000 Mc/s must almost certainly originate within the positive column enclosed by the resonant cavity; using the theory of Rosen (1949) it was calculated that the cylindrical tubes surrounding the positive column on either side of the cavity acted as waveguides beyond cut-off at 3000 Mc/s, for discharge currents less than 70 mA. Under this condition noise generated externally did not enter the cavity. A similar limiting value of the discharge current (70 mA) was deduced from standing wave measurements at 3000 Mc/s with the discharge as load. It is assumed that the additional source of noise in the positive column which is evident at 3000 Mc/s had a spectrum extending to 200 Mc/s.

It may be concluded that there are at least three mechanisms which contribute to the enhanced values of T' at 200 Mc/s shown in Figure 7:

- (i) radiation generated in the vicinity of the electrodes, which may enter the cavity by coaxial transmission;

- (ii) fluctuations in the discharge current at 200 Mc/s which may contribute to the radiation field within the cavity; and
- (iii) generation by the plasma within the cavity of radiation which is in excess of the thermal level.

The experimental results do not permit us to state quantitatively the relative magnitudes of the contributions from these three mechanisms.

VI. ACKNOWLEDGMENTS

It is a pleasure to thank Mr. F. J. Leahy, Dr. J. L. Pawsey, and Mr. L. G. Dobbie for discussions; Mr. F. C. James for his assistance in the preparation of the discharge tubes; and Dr. R. N. Bracewell for a critical reading of the manuscript.

VII. REFERENCES

- CAHN, J. H. (1949).—*Phys. Rev.* **75**: 293.
 CHAPMAN, S., and COWLING, T. G. (1951).—"The Mathematical Theory of Non-uniform Gases." 2nd Ed. (Cambridge Univ. Press.)
 DICKE, R. H. (1946).—*Rev. Sci. Instrum.* **17**: 268.
 DRUYVESTEYN, M. J., and PENNING, F. M. (1940).—*Rev. Mod. Phys.* **12**: 87.
 EASLEY, M. A., and MUMFORD, W. W. (1951).—*J. Appl. Phys.* **22**: 846.
 VON ENGEL, A., and STEENBECK, M. (1934).—"Elektrische Gasentladungen." Vol. 2. (Springer: Berlin.)
 FOWLER, R. H. (1936).—"Statistical Mechanics." 2nd Ed. (Cambridge Univ. Press.)
 JOHNSON, H., and DEREMER, K. R. (1951).—*Proc. Inst. Radio Engrs., N.Y.* **39**: 908.
 KNOL, K. S. (1951).—*Philips Res. Rep.* **6**: 288.
 LABRUM, N. R., and BIGG, E. K. (1952).—*Proc. Phys. Soc. Lond. B* **65**: 356.
 LOVE, A. W. (1948).—*J. Instn. Engrs. Aust.* **20**: 33.
 MIERDEL, G. (1938).—*Wiss. Veröff. Siemens* **17**: 71.
 MUMFORD, W. W. (1949).—*Bell Syst. Tech. J.* **28**: 608.
 NIELSEN, R. A. (1936).—*Phys. Rev.* **50**: 950.
 ROSEN, P. (1949).—*J. Appl. Phys.* **20**: 868.
 SLATER, J. C. (1950).—"Microwave Electronics." (D. van Nostrand & Co. Inc.: New York.)
 SMERD, S. F., and WESTFOLD, K. C. (1949).—*Phil. Mag.* **40**: 831.
 WESTFOLD, K. C. (1950).—*Phil. Mag.* **41**: 509.
 WESTFOLD, K. C. (1953).—*Phil. Mag.* **44**: 711.
 WOOLLEY, R. v. d. R. (1947).—*Aust. J. Sci.* **10** (2): Suppl.

APPENDIX I

By L. W. DAVIES

Radio-frequency Radiation from Non-equilibrium Discharges in Neon

A discussion is given here of the intensity level of the radio-frequency radiation from a plasma in which the electrons are assumed to have a distribution of energies derived by Druyvesteyn (Chapman and Cowling 1951, Sect. 18.71); the distribution applies to a plasma in which there is negligible interaction between electrons. The problem is approached by deriving the effective temperature for radio-frequency radiation of a plasma of infinite optical depth, when the radiation (bremsstrahlung) is emitted and absorbed during free-free encounters of electrons with positive ions.

The equation of transfer of radio-frequency radiation along a ray trajectory passing through an ionized medium may be written (Woolley 1947)

$$\mu_f^2 \frac{d}{ds} \left(\frac{I_f}{\mu_f^2} \right) = \eta_f - K_f I_f, \quad \dots\dots\dots (\text{A1})$$

where I_f is the intensity of radiation in the frequency range $(f, f+df)$ in the direction of the ray, η_f is the volume emissivity, K_f the absorption coefficient, μ_f the refractive index of the medium, and s the distance measured along a ray trajectory.

Upon the introduction of the optical depth τ_f , defined by

$$\tau_f(s) = \int_0^s K_f ds, \quad \dots\dots\dots (\text{A2})$$

the solution of (A1) is

$$\frac{I_f}{\mu_f^2} = \left(\frac{I_f}{\mu_f^2} \right)_0 e^{-\tau_f} + e^{-\tau_f} \int_0^{\tau_f} F_f e^{\tau_f} d\tau_f, \quad \dots\dots\dots (\text{A3})$$

where

$$F_f = \frac{\eta_f}{K_f \mu_f^2} \quad \dots\dots\dots (\text{A4})$$

is the ergiebigkeit of the medium for radiation in the frequency range $(f, f+df)$ (Smerd and Westfold 1949), and the subscript 0 refers to the zero level for τ_f .

If the ionized medium is uniform in its properties, so that the ergiebigkeit F_f is independent of optical depth, it follows from equation (A3) that

$$\frac{I_f}{\mu_f^2} = \left(\frac{I_f}{\mu_f^2} \right)_0 e^{-\tau_f} + F_f (1 - e^{-\tau_f}), \quad \dots\dots\dots (\text{A5})$$

and for a medium of infinite optical depth,

$$\frac{I_f}{\mu_f^2} = F_f. \quad \dots\dots\dots (\text{A6})$$

Hence the intensity of radiation is directly related to the ergiebigkeit for such a medium. In the particular case of a medium in thermodynamic equilibrium at temperature T we have from the Kirchhoff relationship

$$F_f = B_f(T) = \frac{2k}{c^2} f^2 T, \quad \dots\dots\dots (\text{A7})$$

where $B_f(T)$ is the intensity of black radiation at temperature T . Thus in this case

$$I_f = \mu_f^2 \frac{2k}{c^2} f^2 T. \quad \dots\dots\dots (\text{A8})$$

For non-equilibrium conditions we define an effective temperature T' for the emergent radiation by means of the relationship

$$I_f = \mu_f^2 \frac{2k}{c^2} f^2 T'. \quad \dots\dots\dots (\text{A9})$$

Smerd and Westfold (1949) and Westfold (1950) have calculated the *ergiebigkeit* for an ionized medium in equilibrium from the microscopic processes occurring in the plasma; they assumed that the contributions to the r.f. radiation field arose from the bremsstrahlung emitted by electrons during their collisions with positive ions. Making this same assumption, we derive the *ergiebigkeit*, in this case, for an ionized medium in which the electrons have a distribution of energies $\mu(\eta)$ given by Druyvesteyn, namely,

$$\mu(\eta) = B\eta^{\frac{1}{2}}e^{-\alpha\eta^{\frac{1}{2}}/\bar{\eta}}, \quad \dots\dots\dots (\text{A10})$$

where $\alpha = [\Gamma(5/4)/\Gamma(3/4)]^2$, $\bar{\eta}$ is the mean energy, and from the relationship

$$\int_0^\infty \mu(\eta) d\eta = N_1,$$

where N_1 is the electron density, it follows that

$$B = 2N_1\alpha^{3/4}/\{\bar{\eta}^{3/2}\Gamma(3/4)\}.$$

When the motion of the positive ions in the plasma is neglected, the number of collisions per unit time per unit volume between ions and electrons such that the electron energy η lies in the range $(\eta, \eta + d\eta)$ and the collision parameter b lies in $(b, b + db)$ is given by (Fowler 1936, Sect. 19.32)

$$N_2 \cdot 2\pi b \cdot db \cdot \left(\frac{2\eta}{m}\right)^{\frac{1}{2}} \mu(\eta) d\eta,$$

where N_2 is the number-density of positive ions, and m is the mass of the electron.

It follows that the number ν of collisions per unit time made by one electron with the ions of the medium is

$$\nu = \frac{1}{N_1} \int_0^\infty \int_0^\sigma N_2 \cdot 2\pi b \left(\frac{2\eta}{m}\right)^{\frac{1}{2}} \mu(\eta) db d\eta. \quad \dots\dots (\text{A11})$$

In this expression the upper limit σ to the variable b is an effective collision distance for the ions. It has been discussed by Smerd and Westfold (1949) and Westfold (1953).

From Smerd and Westfold's treatment we have

$$K_f = 4\nu f_0^2 / 3ef^2\mu_f, \quad \dots\dots\dots (\text{A12})$$

where

$$f_0^2 = N_1 e^2 / \pi m. \quad \dots\dots\dots (\text{A13})$$

The spontaneous emission per unit volume per unit time is given by

$$4\pi\eta_f \cdot df = \int_0^\infty \int_0^\sigma Q_f \cdot df \cdot N_2 \cdot 2\pi b \left(\frac{2\eta}{m}\right)^{\frac{1}{2}} \mu(\eta) db d\eta, \quad \dots\dots (\text{A14})$$

where $Q_f df$ is the energy radiated in the frequency range $(f, f+df)$ as bremsstrahlung during a collision between electron and positive ion. Its value is given by Westfold (1950)

$$Q_f df = \frac{16e^2}{3c^3} \cdot \frac{2\eta}{m} \cdot \frac{\mu_f}{\{1 + (2\eta b / Ze^2)^2\}}, \dots\dots\dots (\text{A15})$$

where Ze is the charge on the positive ion.

We are now in a position to evaluate the ergiebigkeit for arbitrary electron energy distribution $\mu(\eta)$, and thence the effective temperature T' from equations (A6) and (A9). We shall do so first for the case of a Maxwellian distribution of electron velocities. In this case

$$\mu(\eta) = 2\pi N_1 \eta^{\frac{1}{2}} (\pi k T_e)^{-3/2} \exp(-\eta / k T_e), \dots\dots\dots (\text{A16})$$

whence we have from equations (A11) and (A14):

$$\nu = 2N_2 \sigma^2 (2\pi k T_e / m)^{\frac{1}{2}}, \dots\dots\dots (\text{A17})$$

and

$$\eta_f = \frac{4\sqrt{2}N_1 N_2 Z^2 e^6 \mu_f}{3\sqrt{\pi} c^3 (m k T_e)^{3/2}} \int_0^\infty \left[\ln \left\{ 1 + \frac{4\eta^2 \sigma^2}{Z^2 e^4} \right\} \right] \exp(-\eta / k T_e) d\eta,$$

that is,

$$\eta_f = 8N_1 N_2 Z^2 e^6 \mu_f A_1(2) / \{3mc^3(2\pi m k T_e)^{\frac{1}{2}}\}, \dots\dots\dots (\text{A18})$$

where $A_1(2)$ is the mean value of the slowly varying logarithmic function $\ln\{1 + 4\eta^2 \sigma^2 / Z^2 e^4\}$, and we have made the approximation of replacing this function by its mean value in the integral. Smerd and Westfold give

$$\sigma = \{A_1(2)\}^{\frac{1}{2}} Z e^2 / 2k T_e. \dots\dots\dots (\text{A19})$$

From equations (A4), (A12), (A13), (A16), (A17), and (A18) we find for the ergiebigkeit

$$\begin{aligned} F_f &= \frac{2k}{c^2} f^2 T_e \\ &= B_f(T_e). \end{aligned}$$

Thus the effective noise temperature of a plasma, in which the electrons have a Maxwellian distribution of energies with temperature T_e , is equal to the electron temperature.

We now calculate the ergiebigkeit for the Druyvesteyn distribution, equation (A10).

From equations (A10) and (A11) we have

$$\nu = \pi N_2 \sigma^2 B \bar{\eta}^2 / \{\alpha N_1 (2m)^{\frac{1}{2}}\}. \dots\dots\dots (\text{A20})$$

Similarly from equations (A10), (A14), and (A15),

$$\eta_f = \frac{4BN_2 Z^2 e^6 \mu_f}{3mc^3 (2m)^{\frac{1}{2}}} \int_0^\infty \left[\ln \left\{ 1 + \frac{4\eta^2 \sigma^2}{Z^2 e^4} \right\} \right] \exp(-\alpha \eta^2 / \bar{\eta}^2) d\eta,$$

that is,

$$\eta_f = (2\pi)^{\frac{1}{2}} B N_2 Z^2 e^6 \mu_f A_1(2) \bar{\eta} / \{3mc^3 (m\alpha)^{\frac{1}{2}}\}, \dots\dots\dots (\text{A21})$$

where we have again made the approximation of replacing the slowly varying logarithmic function by its (constant) mean value. The equivalent relationship to equation (A19) is now

$$\sigma = 3\{A_1(2)\}^{\frac{1}{2}} Z e^2 / 4 \bar{\eta}. \quad \text{..... (A22)}$$

From equations (A4), (A12),* (A13), (A19), (A20), and (A21) we have for the ergiebigkeit

$$F_f = 8(\pi\alpha)^{\frac{1}{2}} f^2 \bar{\eta} / 9 c^2. \quad \text{..... (A23)}$$

Inserting the value $\alpha = [\Gamma(5/4)/\Gamma(3/4)]^2$ we find

$$F_f = \frac{2f^2}{c^2} \left[\frac{4\sqrt{\pi}}{9} \frac{\Gamma(5/4)}{\Gamma(3/4)} \right] \bar{\eta},$$

that is,

$$\begin{aligned} F_f &= \frac{2f^2}{c^2} [0.874 k T_e] \quad \text{..... (A24)} \\ &= 0.874 B_f(T_e), \end{aligned}$$

where the electron temperature T_e is related to the mean energy by

$$\bar{\eta} = \frac{3}{2} k T_e.$$

Thus the calculated value of effective noise temperature of the plasma in this case is $0.874 T_e$.

* Dr. Westfold has kindly pointed out that the use here of equation (A12) involves the assumption that the expression for K_f remains unchanged for a Druyvesteyn distribution.

THE CHARACTERISTICS OF TRACKS IN NUCLEAR RESEARCH EMULSIONS

By A. J. HERZ* and G. DAVIS†

[*Manuscript received September 27, 1954*]

Summary

Formulae are derived giving the dependence of the blob, grain, and gap densities, the total gap length, and the mean gap width in the tracks of charged particles as function of the probability of making a silver halide crystal developable. The growth of the grains during the development process is taken into account. The expressions found are suitable for use in discussions of the effects of changes in emulsion properties, and they show clearly the differences between the various methods of determining the rate of energy loss from the track characteristics. There is good agreement with experiment for both G5 and C2 emulsion.

I. INTRODUCTION

In the analysis of tracks of charged particles in nuclear research emulsions, the ionizing power of the particles is estimated from the characteristics of the tracks. Basically, the probability of making a silver halide crystal developable increases with the rate of energy loss by ionization of the particle, and any property of the track which depends on this probability can be used as a measure of the rate of energy loss. The earliest, and most obvious, of the properties thus used was the grain density, i.e. the number of developed grains per unit length of track. However, the measurement of grain density is highly subjective as soon as the grain density is so great that the grains occur in groups rather than as separate units. In addition, the grain density asymptotically approaches a maximum value, so that it is not a sensitive measure of the ionizing power when the track is a dense one.

A large number of techniques has been proposed and used by many workers in efforts to make the estimation of the rate of energy loss more accurate, more objective, and more reliable, and only a few can be mentioned here. For tracks of very low density, Voyvodic (1951) suggested the replacement of the grain density by the density of unresolved grain groups ("blobs"), and this technique has become generally accepted. The blob density is, of course, equal to the density of the gaps between the blobs. Hodgson (1950) first published the method in which the fraction of track occupied by gaps (that is the gap length per unit length of track) is measured. This method is particularly suitable for very dense tracks. The measurement of gap density (the number of gaps

* F. B. S. Falkiner Nuclear Research Laboratory, School of Physics, University of Sydney ; also supported by the Nuclear Research Foundation within the University of Sydney.

† Imperial College of Science and Technology, London, England ; present address : the Technion, Haifa, Israel.

per unit length of track) has been discussed by O'Ceallaigh (1953) and Tennent (1953), and O'Ceallaigh (1953) has suggested the mean gap width as a measure of the ionizing power of the particle, whilst Tennant (1953) proposed the counting of gaps wider than a certain minimum.

Apart from these "visual" methods, a number of photometric ones have been developed and used (see, for example, von Friesen and Kristiansson 1951; Ceccarelli and Zorn 1952).

In general, all the track characteristics which are measured depend on the probability p that a specified silver halide crystal in the path of the particle be rendered developable. In the present paper, we give the relations describing the dependence on p of the various quantities observed in the visual methods. These relations allow one to discuss quantitatively the merits and ranges of usefulness of the different measuring techniques, and they can also be used to predict the effects of changes in emulsion properties.

In general, our approach is similar to that of Della Corte, Ramat, and Ronchi (1953), but it is more rigorous, and we have gone further in deriving the consequences of our model.

II. THE MODEL

In the derivation of the formulae a model is used in which all the fluctuations in the parameters have been neglected.

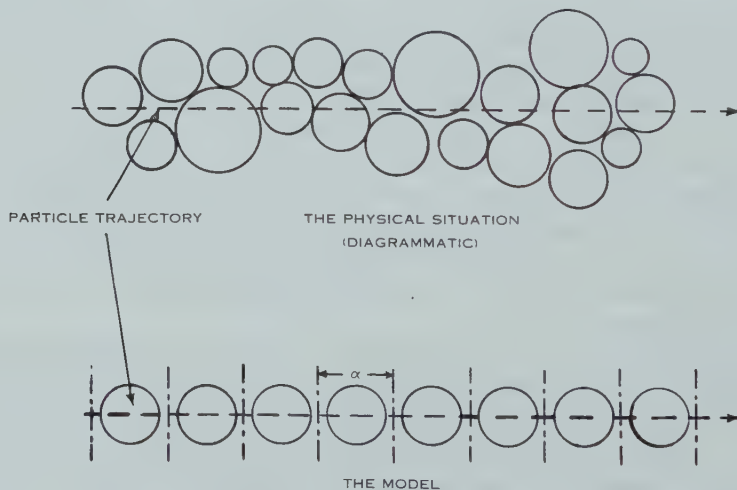


Fig. 1.—Comparison between the physical situation and the model.

It is assumed (see Fig. 1) that the track can be divided up into cells, all of equal length α , each containing one silver halide crystal. The probability that the crystal in any particular cell be developed is p . When the crystal is developed it grows to a size greater than α by a factor γ , so that the developed grain size (the diameter if the grain is spherical) is $\gamma\alpha$. Figure 2 shows the effect of this. It is clear that with this model blobs can never be less than $\gamma\alpha$ in diameter, but, depending on the value of γ , much smaller values of the gap width are possible.

III. DERIVATION OF THE FORMULAE

It is convenient to consider first the formation of the gaps in the tracks, for gaps are simple structures consisting only of a series of undeveloped cells with a developed one at each end.

To start with we find the width of the (in theory) smallest possible gap. From Figure 2 it is seen that the width of this "first order" gap is $\alpha(\Gamma+1-\gamma)$, where Γ denotes the integral part of γ , that is, the nearest integer less than or equal to γ . The width of the next bigger gap, which we call a gap of order $i=2$, is (see Fig. 2) $\alpha(\Gamma+2-\gamma)$, and similarly, the width of an i th order gap will be

$$w_i = \alpha(\Gamma + i - \gamma). \quad \dots\dots\dots (1)$$

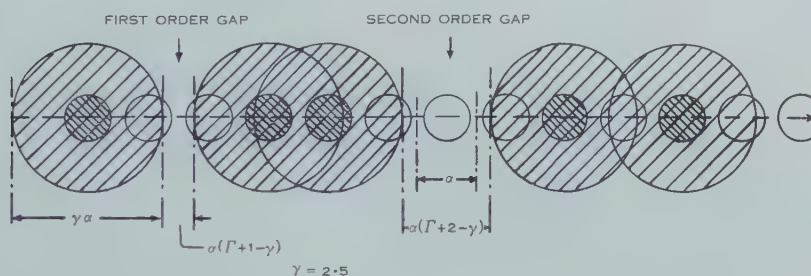


Fig. 2.—Details of the model. Small shaded circles represent developable crystals, large ones are developed grains.

The probability of finding an i th order gap starting in a specified cell of the track is the probability of getting two developable cells with $(\Gamma+i-1)$ undevelopable ones in between: $p^2(1-p)^{\Gamma+i-1}$. Hence, if the number of cells in track length t is T , the expected density of i th order gaps will be

$$\bar{n}_i(t) = Tp^2(1-p)^{\Gamma+i-1}, \quad \dots\dots\dots (2)$$

as long as $w_i \ll t$.

The gap density and the blob density are equal, and they are obtained by summing \bar{n}_i over all values of i :

$$\bar{n}_b = Tp(1-p)^{\Gamma}. \quad \dots\dots\dots (3)$$

The total length of gap in track length t is given by

$$\bar{x} = \sum_{i=1}^{\infty} w_i \bar{n}_i = t(1-p)^{\Gamma} \{1 - p(\gamma - \Gamma)\}. \quad \dots\dots\dots (4)$$

The total length of the blobs in length t will be

$$L_b = t - \bar{x} = t[1 - (1-p)^{\Gamma} \{1 - p(\gamma - \Gamma)\}], \quad \dots\dots\dots (5)$$

and, dividing this by the size of the developed grains, $\gamma\alpha$, we get the grain density as customarily defined:

$$\frac{L_b}{\gamma\alpha} = \frac{T}{\gamma} [1 - (1-p)^{\Gamma} \{1 - p(\gamma - \Gamma)\}]. \quad \dots\dots\dots (6)$$

A further important quantity is the expected value of the gap width, given by the particularly simple expression

$$\bar{w} = \alpha \left(\frac{1}{p} + \Gamma - \gamma \right). \quad \dots\dots\dots (7)$$

Finally, if only gaps of order greater than or equal to s (i.e. of width $\geq \alpha(\Gamma + s - \gamma)$) are counted, we find their density to be

$$\bar{n}_{i \geq s} = Tp(1-p)^{\Gamma+s-1}, \quad \dots\dots\dots (8)$$

and the expected value of their width becomes

$$\bar{w}_{i \geq s} = \alpha \left(\frac{1+p(s-1)}{p} + \Gamma - \gamma \right) = \bar{w} + \alpha(s-1). \quad \dots\dots (9)$$

IV. EXPERIMENTAL TESTING OF THE RELATIONS

Of the adjustable parameters, the size of the developed grains, $\gamma\alpha$, is found by plotting a distribution of blob sizes for tracks with a blob density near the minimum value. In such tracks, most of the grains will be single, that is, grown from single silver halide crystals, and they give a large peak in the distribution. We make the reasonable assumption that the distribution of single-grain sizes is symmetrical, so that its mean is equal to the mode of the total distribution. We thus take the mode as our value of $\gamma\alpha$.

TABLE 1
REPRESENTATIVE RESULTS FOR TRACKS IN G5 EMULSION
 $\alpha = 0.30 \mu$, $\gamma = 2.00$, $t = 100 \mu$

Track No.	Observed			Calculated		
	\bar{w} (μ)	\bar{n} (100)	L_G (μ)	p	\bar{n} (100)	L_G (μ)
1	2.01	36.0	72.5	0.149	36.0	72.4
2	1.64	42.3	69.4	0.164	40.7	66.8
3	1.52	42.8	64.6	0.197	42.3	64.5
4	1.33	46.7	62.0	0.226	45.1	59.9
5	1.19	47.5	57.0	0.250	46.9	56.2
6	0.64	46.2	29.4	0.469	44.1	28.2
7	0.60	39.8	23.7	0.503	41.4	24.7

The cell size α has to be found by trial and error, using a dense track, as the characteristics of dense tracks are most sensitive to changes in the parameters. It is known from electron-microscope measurements that the size of the undeveloped crystals in Ilford G5 emulsion is of the order of 0.3μ (Baroni and Castagnoli 1950; George, personal communication 1952; Pickup 1953; Bradley 1954) and the trial values of α are therefore taken in this region.

Once a value of α has been found which gives the correct characteristics for a dense track in a given plate, it is also satisfactory for all the other tracks in that plate. The values of p are in all cases determined from the observed mean gap width \bar{w} and equation 7). Table 1 gives a representative set of

results for tracks observed in a G5 plate; similar agreement has been obtained in other plates, both of types G5 and C2, but, of course, with different values of γ and α . That the values of α and γ are critical can be seen from the fact that if we assume that $\alpha=0.29\ \mu$ or $\alpha=0.31\ \mu$ for track 7 in Table 1, the calculated blob densities become 45.7 and 73.2 respectively, and the corresponding gap lengths, L_G , are $27.4\ \mu$ and $43.9\ \mu$ in $100\ \mu$ of track. Note that there is a minimum at $\alpha=0.30\ \mu$.

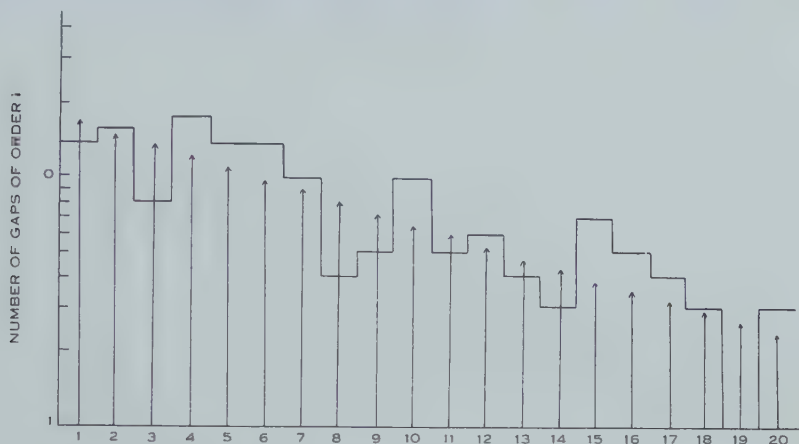


Fig. 3.—Gap size distribution. Histogram, observed; arrows, calculated.

Another way of testing the model is to compare the gap size distribution obtained from equation (2) with the observed one. A representative example is given in Figure 3 where the number of gaps of order i is plotted against i . The calculated values are derived from equation (2), using the values of T , α , and γ which fit all the tracks in the plate. The agreement is very good, especially as the histogram has not been normalized to the calculated values

V. DISCUSSION

Figure 4 shows an example of the calculated variation with p of the various track characteristics in a sample of G5 emulsion. The minimum value of the rate of energy loss gives a blob density of the order of 20–25 blobs per $100\ \mu$ and this corresponds to a value of the probability p of order 0.05. The steep slope of the blob-density curve in this region confirms the well-known fact that for weakly ionizing particles the blob density is a sensitive measure of the rate of energy loss. If the ionization loss of the particle rises above about twice the minimum, however, the blob density becomes quite insensitive to changes in p , and the total and mean gap lengths become preferable. In Figure 5, the density of gaps greater than a given size is plotted against p , showing that this quantity varies rapidly with p for dense tracks, and that it does not have as extensive a plateau at medium values of p as the blob density. If the statistical accuracy is sufficient, therefore, this measure can give good discrimination, as has been found experimentally by Tennent (1953).

How closely our model approaches the real situation depends on the way in which the crystals of the undeveloped emulsion are distributed in space, and on the size distribution of the crystals. We know from the distributions of crystal sizes in Ilford G5 emulsion which have been published (Baroni and Castagnoli 1950; Pickup 1953) that the variations in crystal size are quite small, and it therefore seems reasonable to use a model in which the crystals are assumed to be all of the same size. We are left, then, with the question "How are the crystals distributed in space?", and here our model assumes extreme ordering, with all cell sizes equal, whilst O'Ceallaigh (1953) postulates complete disorder with an exponential distribution of the distances between the crystals.

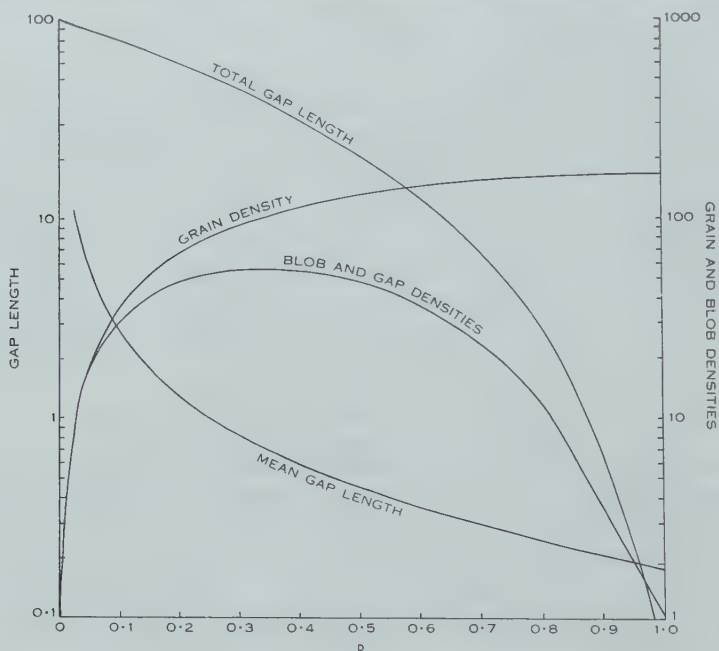


Fig. 4.—Variation of track characteristics with p in a sample of G5 emulsion ($\alpha=0.29 \mu$, $\gamma=2.24$).

Which of these two basic assumptions is nearer to the truth seems to be a question which can only be decided by experiment after the full consequences of the assumption of disorder, made by O'Ceallaigh (1953), have been worked out, allowing one to compare the numerical predictions of both theories. Meanwhile we feel that the somewhat amazing agreement between the experimental data and the detailed numerical results derived from our model is sufficient to show that it must be adequate in most cases of practical importance.

There is, however, one point where there is considerable disagreement between our model and O'Ceallaigh's. O'Ceallaigh's gap width distribution is a continuous exponential function, whilst ours (equation (2)) is a discrete set of values decreasing according to a power law. Now it is easily shown (O'Ceallaigh 1953) that a continuous exponential distribution would make the mean gap width independent of the degree of grain growth (γ in our notation), so that \bar{w}

could be used as a measure of the probability p unaffected by physical development. On the other hand, if the distribution is as given in equation (2) then there is a dependence on the fractional part of γ , $\Gamma - \gamma$, and this is particularly important when p is large, that is, for dense tracks. In view of these uncertainties we feel that the claim (O'Ceallaigh 1953; Johnston and O'Ceallaigh 1954) that the mean gap width is independent of the amount of grain growth by physical development requires experimental proof before it can be accepted.

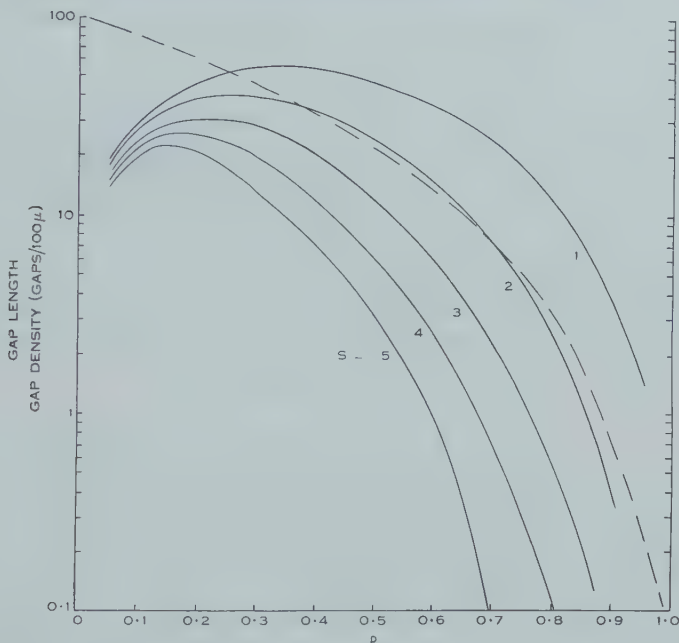


Fig. 5.—Density of gaps in order $\geq s$ in G5 emulsion ($\alpha=0.29 \mu$, $\gamma=2.24$). Broken line: total gap length per 100 μ .

VI. ACKNOWLEDGMENTS

The authors would like to express their thanks to Dr. E. P. George for a number of helpful discussions; they are greatly indebted to Mr. C. Waller for much useful advice on the properties of nuclear research emulsions.

VII. REFERENCES

- BARONI, G., and CASTAGNOLI, C. (1950).—*Nuovo Cim.* **7**: 364.
 BRADLEY, D. E. (1954).—*Brit. J. Appl. Phys.* **5**: 96.
 CECCARELLI, M., and ZORN, G. T. (1952).—*Phil. Mag.* **43**: 356.
 DELLA CORTE, M., RAMAT, M., and RONCHI, L. (1953).—*Nuovo Cim.* **10**: 509.
 VON FRIESEN, S., and KRISTIANSSON, K. (1951).—*Nature* **168**: 377.
 HODGSON, P. E. (1950).—*Phil. Mag.* **41**: 725.
 JOHNSTON, R. H. W., and O'CEALLAIGH, C. (1954).—*Phil. Mag.* **45**: 424.
 O'CEALLAIGH, C. (1953).—Rep. Cosmic Ray Congress, Bagnères-de-Bigorre, p. 73.
 PICKUP, E. (1953).—*Canad. J. Phys.* **31**: 898.
 TENNENT, R. M. (1953).—Ph.D. Thesis, University of London.
 VOJVODIC, L. (1951).—Proc. Bristol Conf. on V Particles and Heavy Mesons, p. 16.

OBSERVATIONS ON THE PENETRATING COMPONENT OF EXTENSIVE AIR SHOWERS

By V. C. OFFICER* and P. J. ECCLES*

[*Manuscript received December 1, 1954*]

Summary

The variation in the composition of the penetrating component of extensive air showers at sea-level has been investigated as a function of shower density. The proportion of N-particles in the penetrating component was observed to decrease as the shower density decreased. It is suggested that this behaviour is due to the filtering action of the atmosphere on the more absorbable particles in the nuclear cascade which has passed maximum development.

The zenith angle distribution of the penetrating μ -mesons in air showers has been found to be of the form $\cos^2 \theta$ which is consistent with that expected for shower axes, and gives no evidence for the broader distribution reported at mountain altitudes.

An anomalous deficiency of penetrating μ -mesons with zenith angles less than 5° has been found in air showers of median density 10 m^{-2} .

I. INTRODUCTION

The data on the penetrating component of extensive air showers reported in this paper were collected during an experiment in which the penetrating component was searched for delayed particles (Officer and Eccles 1954). Analysis of hodoscope pictures taken under 15 cm of lead gave information on the composition of the extensive shower penetrating component, the zenith angle distribution of the penetrating μ -mesons, and the dependence of these on shower density.

Since the early work on the penetrating particles in air showers discussed by Broadbent and Jánosy (1948), there have been many investigations on this subject. These include researches by Cocconi, Cocconi Tongiorgi, and Greisen (1949), Mitra and Rosser (1949), Ise and Fretter (1949), Cocconi and Cocconi Tongiorgi (1950), McCusker (1950), Greisen, Walker, and Walker (1950), Sitte (1950), McCusker and Millar (1951), Sitte (1952), Hodson (1953), and Kasnitz and Sitte (1954). There is general agreement that at sea-level about 2 per cent. of shower particles can penetrate 20 cm of lead, and there seems to be a slow increase in this proportion as the shower density decreases. Most of the penetrating particles are μ -mesons and the remainder called the N-component, are nuclear interacting particles such as neutrons, protons, and π -mesons. However, the composition of the penetrating component and its abundance as a function of distance from the shower core and of shower age, have not been fully investigated, as shown by the following apparently conflicting results. Several workers (Sitte 1950, 1952 ; Hodson 1953 ; Piccioni and Cool 1953) have reported

* Physics Department, University of Melbourne.

proportionality between the penetrating particle and electron densities. Recently Kasnitz and Sitte (1954) working at an altitude of 3260 m have found the abundance of the penetrating component to increase with shower age, due mainly to an increase in the proportion of N-component. The results of the present experiment carried out at sea-level show that there is a marked decrease in the proportion of N-component as shower density decreases.

The zenith angle distribution of air shower axes has usually been studied by examining the inclination of approximately parallel electron tracks in cloud chamber pictures. The direction of the penetrating μ -mesons in air showers should be close to that of the shower axis, but Brown and McKay (1949) and Sitte (1950) working at mountain altitudes found broader zenith angle distributions for the penetrating particles than for the shower electrons at the same altitudes. The present distribution for the penetrating μ -mesons is in agreement with that expected for the shower axes as calculated from the altitude dependence of air showers by Kraybill (1954).

The zenith angle distribution for the penetrating μ -mesons has been examined for three ranges of shower density. An anomalous deficiency of zenith angles less than 5° compared with the numbers expected for a cosine power law, has been found in the group with a median density of 10 particles/m².

II. EXPERIMENTAL ARRANGEMENT

The showers were detected by a threefold coincidence arrangement consisting of trays of Geiger counters at the vertices of a 2 m triangle. One tray, 359 cm² in area, was shielded with 15 cm of lead and the other two trays, 450 cm² and 583 cm² in area, were unshielded. The resolving time was approximately 1 μ sec. Immediately above and below the shielded tray were two trays of counters connected to a hodoscope, and below these were arranged alternately, three 5 cm lead layers and three trays of hodoscope counters. The ends and sides of the array were shielded with 10 cm of lead. The remaining five channels of the 50 channel hodoscope were connected to a simple form of shower density detector consisting of one sensitive area of 500 cm², three of 90 cm², and one of 15 cm². The apparatus is described in detail by Officer and Eccles (1954).

III. COMPOSITION OF THE PENETRATING COMPONENT

(a) *Classification of the Hodoscope Pictures*

The hodoscope pictures of shower components penetrating 15 cm of lead were divided into four main classes, μ -mesons, N-component with range $R > 25$ cm of lead, N-component and electronic event mixture with $15 < R < 25$ cm of lead, and a class containing photons and oblique particles. The criteria used in classifying the pictures were as follows.

(i) *Samples of Penetrating Component containing one or more μ -Mesons.*—Most of the pictures in this class showed the track of a single penetrating non-interacting particle. The discharge of two or several adjacent counters under one of the 5 cm layers of lead was allowed for inclusion in the class, but the frequency of these occurrences was consistent with the production of knock-on electrons and knock-on showers by μ -mesons. There were 240 of these single

particle pictures and an additional small group of 11 pictures that could have been very oblique particles in the top cavity. A group of 71 pictures, showing an otherwise single track accompanied by an apparent knock-on event in the top lead cavity immediately below the upper 15 cm lead shield, was included. The frequency of these events was such that at least half of them must have represented coincidences in the top cavity between μ -mesons and other particles from the air showers, probably of electronic nature. Also included was a group of 31 pictures showing parallel penetrating tracks. One picture showed four tracks, two showed three tracks, and the rest had only two parallel tracks. Two of these particles were N-particles but the rest were considered to be μ -mesons.

(ii) *Samples containing N-component.*—Penetration of at least 25 cm of lead was required. The pictures usually showed several counters discharged in most of the layers traversed, but any particle which gave rise to an interaction with penetrating secondaries was included. In addition to 136 events of this type 32 events in which most of the counters in each layer were discharged were included. In some cases all 45 shielded hodoscope channels registered a hit. Samples of penetrating component in this class thus contained one or more N-particles and an unknown number of μ -mesons.

(iii) *Mixed Samples of N-component and Electronic Events.*—This class contained 31 events with ranges between 20 and 25 cm of lead or with a lower limit of 20 cm due to escape from the side of the array. They resembled N-component pictures but identification was less certain than for the longer-range N-component. Also included were 58 pictures showing showers or groups of two or more particles in the top cavity only. As the thickness of lead above this cavity was only 15 cm they could have been produced by a mixture of electronic events and short range N-component.

(iv) *Oblique Particles and Photons.*—Pictures in this class of 143 usually showed no shielded hodoscope counters discharged at all. A considerable number had one counter discharged in the top cavity only, and a few had one or two counters discharged anywhere in the array. These events could not have been accidental coincidences between a background count in the shielded timing tray and a shower striking the two unshielded trays, as these were calculated to have negligible frequency. Also the time interval distribution for this class would have been spread uniformly over the 1 μ sec resolving time if they had been accidental coincidences, but it was sharply peaked (Officer and Eccles 1954). These events were considered to be a mixture of photons and very oblique particles. The photons could have come from burnt-out cascades in the 15 cm of top lead, and the oblique particles must have come through the 10 cm side shielding. If photons were responsible the abundance of the events should have increased with shower density since high energy electrons are more abundant at high densities, but in fact it decreased. This suggests that oblique penetrating particles in low density showers were mainly responsible, and the suggestion fits in with the explanation of a few significant time lags in this group reported earlier (Officer and Eccles 1954).

The classification of the penetrating component samples is summarized in Table 1. In order to obtain the percentage composition of the penetrating

component several corrections to these results would have to be made. As we are at present more interested in the variation of the composition with shower density, these corrections will not be made in full. Firstly, a correction should be made for non-recognition of charged N-particles which do not interact in the thickness of lead traversed. As the 25–30 cm of lead is about two mean free paths thick for N-particles the correction will not be large. Secondly, only those neutrons that interact in the 15 cm upper lead shield will be detected. Greisen, Walker, and Walker (1950) give 1.5 as the ratio of charged to neutral particles in the air shower N-component at 4260 m. Thirdly, the proportion of μ -mesons in the samples containing one or more N-particles is unknown. With large enough samples of penetrating component an N-particle would be included almost invariably, and the present method of classification would yield 100 per cent. N-component. A correction is made for this effect in Section III (b), where it is found to be of importance for only the highest shower density classes recorded.

TABLE 1
CLASSIFICATION OF PENETRATING SAMPLES

Class	No. of Samples	Percentage of Samples
One or more μ -mesons	353	45.8 ± 2.4
N-component (one or more N-particles and unknown μ -meson content)	168	21.8 ± 1.7
N-component and electronic event mixture	89	11.6 ± 1.2
Oblique particles and photons	143	18.6 ± 1.6
Miscellaneous	17	2.2 ± 0.5
Total	770*	

* The total number of detected showers was 782, but in a few cases there was no hodoscope record or the timing record was obscured by a ghost image.

(b) *Variation of the Composition of the Penetrating Component with Shower Density*

Each class of penetrating component samples was analysed according to the combination of large, medium, and small density detector counters discharged. The result of this analysis is shown in Table 2. Rarely discharged density detector combinations have been omitted. Calculation showed that accidental coincidences between background counts in the density detector counters and shower master pulses were insignificant compared with statistical errors in the counter combination totals, in spite of the fact that the resolving time was 20 μ sec.

In order to obtain the median shower densities corresponding to the various density detector counter combinations, several numerical integrations were performed. Firstly, the fraction ε of shower particles able to penetrate 15 cm of lead was found from

$$N = \int_0^{\infty} K \Delta^{-\gamma} (1 - e^{-s_1 \Delta}) (1 - e^{-s_2 \Delta \varepsilon}) (1 - e^{-s_3 \Delta}) d\Delta \quad \text{hr}^{-1},$$

where the observed shower rate $N = 0.50_6 \pm 0.01_8 \text{ hr}^{-1}$, the values of K and γ were taken from Singer (1951), S_1 and S_3 are the areas of the unshielded trays and S_2 is the area of the shielded tray. K and γ were assumed to be independent of shower density Δ . ε was found to be 2.8_6 per cent. and can be compared with 2.60 per cent. found by Cocconi, Cocconi Tongiorgi, and Greisen (1949) under 17.8 cm of lead at an altitude of 260 m . The present higher value could be due to some electronic component penetrating 15 cm of lead.

TABLE 2
ANALYSIS OF SAMPLE CLASSES ACCORDING TO DENSITY DETECTOR COUNTER COMBINATIONS
DISCHARGED

	Counter Combination					
	0	L	LM	L2M	L3M	L3MS
μ -Mesons	65	64	64	41	38	21
N-component	11	17	29	29	25	29
N-component and electronic event mixture	13	8	17	10	12	8
Oblique particles and photons.. ..	30	24	19	21	10	6
Miscellaneous	2	4	5	2	3	—
Totals	121	117	134	106	88	64

Secondly, the frequency of discharge of each counter combination by detected showers was calculated as a function of shower density. The frequency of discharge of a given combination at a density Δ is given by the product of the integrand in the expression for N and the probability of discharging that combination in a shower of density Δ . The probability of discharging the combination L3M is, for example,

$$(1 - e^{-S_L \Delta})(1 - e^{-S_M \Delta})^3 e^{-S_S \Delta},$$

where S_L , S_M , and S_S are the areas of the large, medium, and small density detector counters. For each combination the median density was found as the bisector of the area under the frequency of discharge versus density curve.

It is known that both γ and ε vary slowly with density (Cocconi and Cocconi-Tongiorgi 1949; Cocconi, Cocconi Tongiorgi, and Greisen 1949). The median densities calculated as described will therefore be in error, but this will not upset the main conclusions of this experiment as the exact values of median density are not important. However, these uncertainties do make it difficult to express the results as the variation of the absolute abundance of the various penetrating constituents with density. The variation with density of the percentage composition of the penetrating component is therefore presented.

The data of Table 2 concerning penetrating samples are shown in Figure 1 as percentage composition plotted against median shower density. The increase in the proportion of N-component with increasing density is striking, but part of this rise is due to the inclusion of μ -mesons in the N-component samples.

If it is assumed that μ -mesons and N-particles are both distributed at random in the plane of the detector a simple correction can be applied. The assumption of random mixing of the particles should be valid except for those produced close to the apparatus. If ε_1 and ε_2 are the absolute abundances of

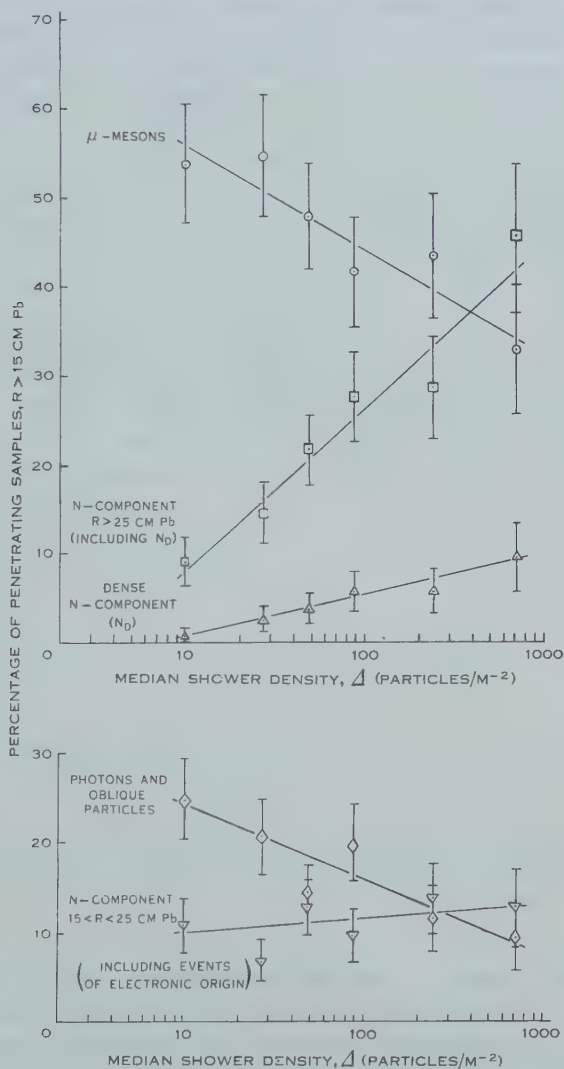


Fig. 1.—Composition of the penetrating component (uncorrected data).

μ -mesons and N-particles at a shower density Δ , the fraction β of penetrating samples observed to contain one or more N-particles will be

$$\beta = \frac{1 - e^{-S\Delta\varepsilon_2}}{1 - e^{-S\Delta(\varepsilon_1 + \varepsilon_2)}},$$

where S is 1200 cm², the area of the hodoscope trays. With the total abundance of penetrating component assumed independent of Δ and given the value 2.8_6

per cent. deduced earlier, $\varepsilon_1 + \varepsilon_2$ can be found from Figure 1 for each Δ and corrected values of ε_1 and ε_2 computed. Besides correcting for μ -mesons included in the N-component samples this procedure also corrects for the fact that the μ -meson samples may contain one or more particles. Figure 2 shows the corrected results for μ -mesons and N-component only. The increase in the proportion of N-component with increasing density is reduced but it is still marked, being a factor of ≈ 3 in the density range 10–710 m^{-2} .

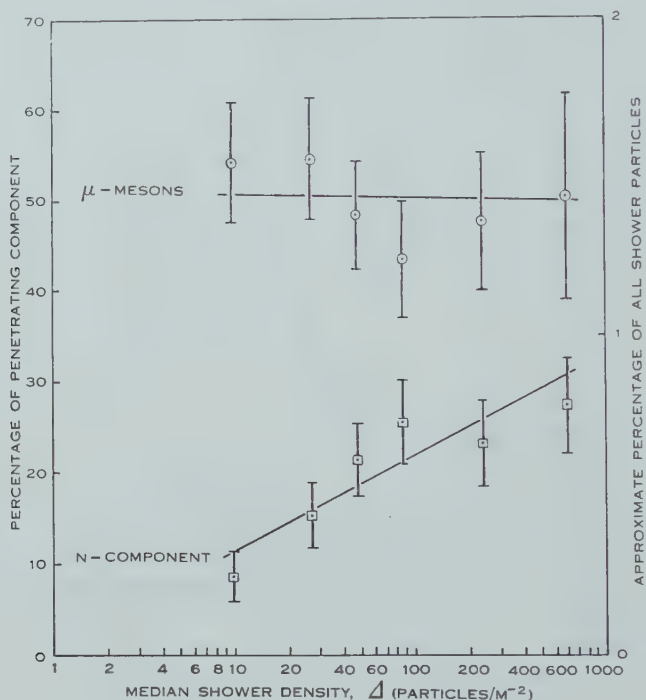


Fig. 2.—Variation with shower density of the percentage of μ -mesons and N-particles in the air shower component penetrating 15 cm of lead (corrected data).

(c) Discussion

As the shower density is reduced, contributions to the counting rate of the present apparatus will come from showers of greater age, and also from points at greater distances from the core in showers that are not necessarily older. It can be shown (Ise and Fretter 1949 ; Singer 1951) that any shower arrangement which is triggered predominantly by showers in a narrow density range will detect mostly showers whose cores strike nearby. This should apply to the present arrangement which includes a shielded tray, unless there is a marked increase in the abundance of penetrating component as distance from the core increases. As other workers have found no evidence for such an increase in the present density range, it is plausible to consider that in Figure 2 decreasing shower density means decreasing shower size (i.e. total number of particles) and thus increasing shower age.

This brings the present results into apparent conflict with those of Kasnitz and Sitte (1954), who found when working at an altitude of 3260 m that the proportion of N-component increased slowly with shower age. They deduced that the nuclear cascade reaches maximum development somewhat later than the electron cascade. They made special efforts to detect low energy N-component, and they suggested that this could continue to multiply after it had ceased to feed energy into the electron cascade via the production of π^0 -mesons. At sea-level, however, both cascades will have passed maximum development and it is probable that we are observing the progressive filtering of the products by the atmosphere as shower age increases. If the process is continued to completion, the low density showers composed chiefly of μ -mesons observed by McCusker and Millar (1951) should be obtained by the absorption of all the less penetrating particles. The present results probably represent an intermediate stage in the process.

IV. THE ZENITH ANGLE DISTRIBUTION OF THE PENETRATING μ -MESONS

The zenith angles projected on a vertical plane perpendicular to the hodoscope counters were measured for the single penetrating non-interacting particles in the air showers. The projected zenith angle could usually be determined from the hodoscope record to $\pm 3^\circ$ and occasionally much more accurately when

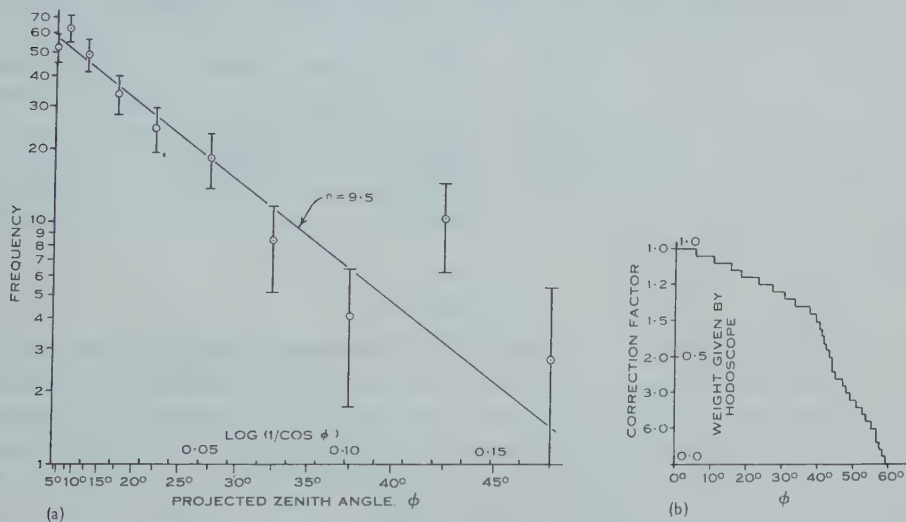


Fig. 3 (a).—Zenith angle distribution of single penetrating μ -mesons in air showers, corrected for geometrical bias of the hodoscope array, but not corrected to a spherical detector.

Fig. 3 (b).—Hodoscope zenith angle sensitivity.

a track passed between two counters in a layer. A few tracks which passed through the top lead cavity only had errors in their zenith angles up to $\pm 5^\circ$, but statistical errors provided the main uncertainties in the distribution.

Figure 3 (a) shows a logarithmic plot of the data corrected for a geometrical bias of the hodoscope array favouring small projected zenith angles. In Figure

3 (b) the weight given by the hodoscope array to each angle is shown. This weight distribution was obtained by the summation of the angular sensitivities, found from the geometry, for tracks passing through each of the small nickel timing counters, a hit on the timing tray being essential for the triggering of the apparatus. The errors shown in Figure 3 (a) are those of statistical origin only. It can be seen that a straight line representing a $\cos^{9.5} \theta$ law is consistent with the corrected data.

As shown by Brown, McKay, and Palmatier (1949) a zenith angle distribution which can be represented tolerably well by a $\cos^n \theta$ law gives rise to a projected angle distribution which is still a cosine power law of the same power but with a new multiplying constant depending on n . Consequently it can be stated that the results are consistent with a zenith angle distribution of the form $\cos^{9.5} \theta$. This applies to showers detected by event-selecting trays which can be represented approximately by three unequal horizontal areas arranged at the vertices of a 2 m triangle and discharged by one or more hits on each. The effect of 15 cm of lead over one tray and the selection of single penetrating particles through it by means of the hodoscope can be regarded as an approximate 100-fold reduction in its area.

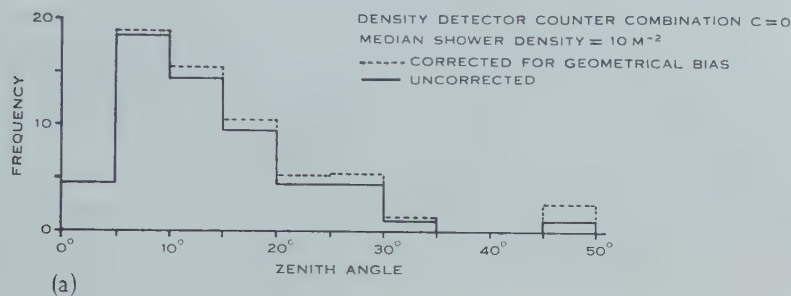
It is usual to state the zenith angle distribution with reference to a spherical detector, and a theorem given by Cocconi, Loverdo, and Tongiorgi (1946) enables this to be done. They showed that, for shower trays of equal area connected in coincidence, reduction of the tray areas by a factor f results in a shower rate reduced by a factor $f^{\gamma-1}$, where $-\gamma$ is the power of the differential shower density distribution. The theorem can be shown to hold also in the present case where trays of unequal area all have the same effective area reduction factor, $\cos \theta$, as the zenith angle is increased. The zenith angle distribution that would be observed with a spherical detector is thus of the form $\cos^8 \theta$ since $\gamma=2.5$.

(a) Discussion

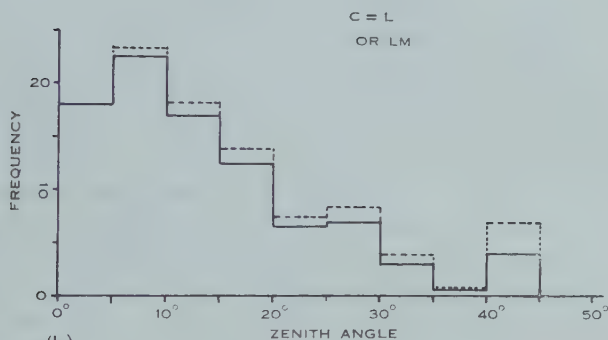
It is to be expected that most of the penetrating μ -mesons observed in extensive air showers at sea-level will travel close to the direction of the shower axis. There will be little scattering of these mesons on account of their high average energy, in the vicinity of several kMeV, and most of them will have been produced at high altitudes such as 5–20 km. Since they are detected amongst the electrons within tens of metres from the shower core, few of them can have large angular deviations from the direction of the shower axis. Kraybill (1954) has calculated the zenith angle distribution of shower axes at sea-level using experimental data on the altitude dependence of air showers. He obtained a cosine power law with the power $n=7.5$, and his result agrees with direct measurements on the electrons at sea-level by Deutschmann (1947). The present value of $n=8$ is consistent with these results.

However, two previous determinations of the zenith angle distribution of penetrating particles in air showers have been carried out at mountain altitudes, and these are not in agreement with the shower axis distributions expected at those altitudes. Brown and McKay (1949) working at an altitude of 3260 m

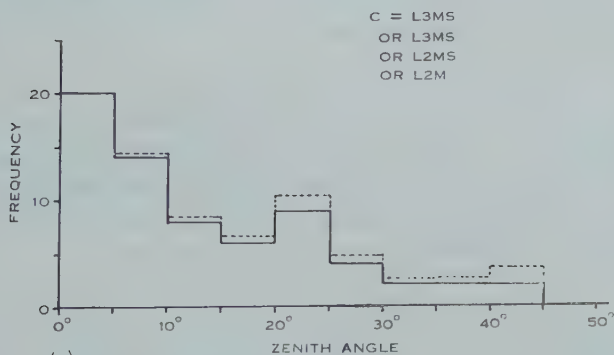
found $n=3$ for all penetrating particles under 15 cm of lead, except those which appeared to have been generated in the lead shield. They found $n=5$ for the electron component and Kraybill (1954) has calculated $n=5.5$ at 3500 m for



(a)



(b)



(c)

Fig. 4.—Penetrating μ -meson zenith angle histograms for three ranges of shower density.

the shower axes. They suspected that local production of penetrating particles in the lead shield may have broadened the distribution. Sitte (1950), who worked at the same altitude, found the zenith angle distribution for penetrating non-interacting particles only. He found $n=3.8$.

The present results do not agree readily with a value of n less than the sea-level shower axis value of 7.5. This is probably accounted for by the present method of observation being insensitive to penetrating particles produced in the lead shielding. Only tracks which appeared single under 15 cm of lead were used, and they frequently penetrated one to three additional 5 cm layers, always without noticeable scattering. It is also possible that at the mountain altitudes, where the nuclear cascade is still developing, there are more wide angle penetrating particles arising from local production in the air than would be expected at sea-level.

It should be noted that Bassi, Clark, and Rossi (1953) have recently obtained a value of n that does not agree with Kraybill's calculated result. They found $n=15 \pm 1.2$ by a new method in which the inclination of shower fronts was found by timing methods.

(b) Dependence of the Zenith Angle Distribution on Shower Density

Figure 4 shows the zenith angle observations grouped according to shower density. The combinations of density detector counters discharged are: zero in Figure 4 (a); L and LM in Figure 4 (b); and L2M, L2MS, L3M, and L3MS in Figure 4 (c). The main feature of interest is the small number of observations falling in the $0-5^\circ$ class in the lowest density group (Fig. 4 (a)). This group has a median density of 10 m^{-2} . The difference between this class and its neighbour of 14 ± 4.8 is 2.9 times its standard error and could be regarded as significant, but more experimental evidence would be desirable. Two other teams, Cresti, Loria, and Zago (1953) and Hazen, Williams, and Randall (1954), have investigated the zenith angle distribution of shower electrons for different densities or shower sizes. They have not reported an anomaly of this kind, but their observations were not restricted to such a narrow range of low densities. Hazen and co-workers were not able to measure zenith angles for shower densities $\leq 200 \text{ m}^{-2}$ because the spread in the directions of the individual electron tracks observed in the cloud chamber was too great.

V. μ -MESON RESIDUAL RANGE DATA

Only 32 mesons appeared to stop in the three 5 cm layers of lead arranged below the 15 cm upper layer. After corrections for particles escaping at the ends and sides of the array, the data were consistent with a uniform distribution of residual ranges in the interval 170–239 g/cm² of lead, as found by Sitte (1950) at an altitude of 3260 m in the interval 280–680 g/cm² of lead.

VI. CONCLUSIONS

The observed decrease in the proportion of N-particles in the penetrating component of air showers at sea-level as shower density decreases is probably the result of progressive filtering of the products of the almost spent nuclear cascade by the atmosphere, leaving the less strongly interacting particles.

The zenith angle distribution of penetrating μ -mesons in air showers at sea-level is consistent with that expected for the shower axes, and gives no evidence of a broader distribution as found at mountain altitudes by other workers.

VII. ACKNOWLEDGMENT

The authors wish to thank Professor L. H. Martin for his interest in this work.

VIII. REFERENCES

- BASSI, P., CLARK, G., and ROSSI, B. (1953).—*Phys. Rev.* **92** : 441.
BROADBENT, D., and JÁNOSSY, L. (1948).—*Proc. Roy. Soc. A* **192** : 364.
BROWN, W. W., and MCKAY, A. S. (1949).—*Phys. Rev.* **76** : 1034.
BROWN, W. W., MCKAY, A. S., and PALMATIER, E. D. (1949).—*Phys. Rev.* **76** : 506.
COCCONI, G., and COCCONI TONGIORGI, V. (1949).—*Phys. Rev.* **75** : 1058.
COCCONI, G., and COCCONI TONGIORGI, V. (1950).—*Phys. Rev.* **79** : 730.
COCCONI, G., COCCONI TONGIORGI, V., and GREISEN, K. (1949).—*Phys. Rev.* **75** : 1063.
COCCONI, G., LOVERDO, A., and TONGIORGI, V. (1946).—*Phys. Rev.* **70** : 841.
CRESTI, M., LORIA, A., and ZAGO, G. (1953).—*Nuovo Cim.* **10** : 779.
DEUTSCHMANN, M. (1947).—*Z. Naturf.* **2** : 61.
GREISEN, K., WALKER, W. D., and WALKER, S. P. (1950).—*Phys. Rev.* **80** : 535.
HAZEN, W. E., WILLIAMS, R. W., and RANDALL, C. A. (1954).—*Phys. Rev.* **93** : 578.
HODSON, A. L. (1953).—*Proc. Phys. Soc. Lond. A* **66** : 65.
ISE, J., Jr., and FRETTER, W. B. (1949).—*Phys. Rev.* **76** : 932.
KASNITZ, H. L., and SITTE, K. (1954).—*Phys. Rev.* **94** : 977.
KRAYBILL, H. L. (1954).—*Phys. Rev.* **93** : 1362.
McCUSKER C. B. A. (1950).—*Proc. Phys. Soc. Lond. A* **63** : 1240.
McCUSKER, C. B. A., and MILLAR, D. D. (1951).—*Proc. Phys. Soc. Lond. A* **64** : 915.
MITRA, S. M., and ROSSER, W. G. V. (1949).—*Proc. Phys. Soc. Lond. A* **62** : 364.
OFFICER, V. C., and ECCLES, P. J. (1954).—*Aust. J. Phys.* **7** : 410.
PICCIONI, O., and COOL, R. L. (1953).—*Phys. Rev.* **91** : 433.
SINGER, S. F. (1951).—*Phys. Rev.* **81** : 579.
SITTE, K. (1950).—*Phys. Rev.* **78** : 721.
SITTE, K. (1952).—*Phys. Rev.* **87** : 351.

RADIO ECHO OBSERVATIONS OF METEORS IN THE SOUTHERN HEMISPHERE

By A. A. WEISS*

[*Manuscript received September 20, 1954*]

Summary

The results of a radio survey of meteor activity at Adelaide are presented. The radiant and activities of six major meteor showers (Geminids, day-time Arietids, ζ -Perseids, δ -Aquarids, Corona Australis, Orionids) have been measured by methods which are described, and the mass distributions in three of these showers are discussed. Seasonal and diurnal variations in the background activity of sporadic meteors are examined in relation to the radiation patterns of the aerial systems. Height distributions for meteors of three showers (Geminids, day-time Arietids, ζ -Perseids) are given. Diurnal variations in the height distribution of sporadic meteors do not conform to those expected from the motion of the apex of the Earth's way.

I. INTRODUCTION

The successful application, at the Jodrell Bank Experimental Station of the University of Manchester, of radio echo techniques to the continuous monitoring of meteor activity in the northern hemisphere, prompted the initiation of a complementary survey in the southern hemisphere. Up to the time of commencement of this survey lists of southern hemisphere visual observations had been published by McIntosh and by Hoffmeister, and McIntosh (1935) had compiled "An Index to Southern Meteor Showers" which lists 320 radiant visible at mid-southern latitudes. These visual observations have since been supplemented by radio echo observations on the δ -Aquarid shower by Hawkins and Almond (1952) and by Lindblad (1952).

Although it is certain that all major southern night-time showers have been detected by the visual workers, the cover in the months September to March is not altogether satisfactory (McIntosh 1935). The chief objects of the southern hemisphere survey were therefore to determine radiant and activities of the known night-time showers, to search for day-time showers, and to measure the hourly rate of sporadic meteors.

This paper presents results obtained from June 1952 to December 1953. The survey was not completed during this period, portions of the first half of the year not being covered. The radiant and activities of the major meteor showers visible from Adelaide have been measured using two different radio equipments. No major day-time showers have been discovered. Seasonal and diurnal variations in the rate of sporadic meteors have been determined and an attempt is made to relate these variations to the polar diagrams of the aerial systems of the respective equipments. The heights of occurrence of a large

* Department of Physics, University of Adelaide.

number of echoes from sporadic meteors have been examined for seasonal and diurnal variations.

The records also furnish the rate of decay of meteor echoes. As the decay rate is of physical rather than astronomical importance, this topic is reserved for a subsequent paper.

II. THE EQUIPMENT

Two different types of equipment have been used at Adelaide, lat. $-34^{\circ} 45'$. Of these, one (the radiant equipment) is designed for the measurement of shower radiants and activity, whilst the other (the wind equipment) is intended primarily for measurement of upper atmosphere winds using meteor trails as drifting test bodies, and the astronomical information is obtained incidentally.

The radiant equipment is essentially similar to the radio echo apparatus described by Aspinall, Clegg, and Hawkins (1951). In the installation at Adelaide two identical beamed aerial arrays are used, directed at low elevation along azimuths 35.5° S. of E. and 14° N. of E. Each array is adapted for common transmission and reception by means of a TR device. The transmitter delivers into each aerial about 3 kW peak power in double pulses 30 μ sec wide and 150 μ sec apart, at a pulse repetition frequency of 47 c/s and at a frequency of 67 Mc/s. The minimum signal detectable at the receiver input is 10^{-13} W. Amplified echoes from each aerial are fed to an intensity modulated cathode-ray tube and are photographed on film moving in a direction perpendicular to the time-base sweep. Range markers at intervals of 200 km, and time signals at 10-min intervals, are placed on the film at the same time. Film speed is 12 cm/hr.

Each aerial array consists of six Yagi aerials mounted at horizontal distances 1.2λ apart and at a height of 1.5λ above ground. The beams are directed at elevation 9° and have half-power widths $\pm 3.5^{\circ}$ in azimuth and $\pm 4.5^{\circ}$ in elevation.

The time of occurrence and the range of each echo detected with the radiant equipment are measured.

The wind equipment was constructed by and has been fully described by Robertson, Liddy, and Elford (1953), but the following equipment parameters are relevant here. The transmitter delivers 250 W c.w. at a frequency of 26.8 Mc/s, and the minimum detectable signal at the receiver is 3×10^{-14} W. The overall gain of the aerial system is 25. The direction of maximum gain points to the zenith, the first zero occurs at a zenith angle of about 45° , and there is a further minor lobe of low gain at low elevation.

The time of occurrence of each echo detected by the wind equipment is measured. In addition, the direction of the reflection point and the slant range are measured for selected echoes. The remaining echo characteristics measurable with this equipment, namely, line-of-sight drift velocity and rate of decay of the meteor trail, have no astronomical significance.

For echo counting purposes, the sensitivities of the respective equipments are continually monitored, either by metering or by injection of test signals. Despite the double pulses of the radiant equipment and the insertion in the video stages of the receiver of a discriminator unit of the type used by Davies and

Ellyett (1949), interference by noise impulses remains a problem. This form of interference is eliminated by the narrow bandwidth and high film speed of the wind equipment.

Using the equipment parameters listed above, and the scattering formula of Lovell and Clegg (1948), the line density of electrons in the meteor trail detected at limiting sensitivity on the axis of the aerial beam is readily found. These limiting sensitivities are: radiant equipment, 3×10^{10} electrons/cm; wind equipment, 1×10^{11} electrons/cm. In other directions the limiting line densities are of course higher.

III. THE DETERMINATION OF METEOR SHOWER RADIANTS

The methods of radiant determination applicable to either equipment rest on the property of the meteor trail of specular reflection (Hey and Stewart 1947; Lovell, Banwell, and Clegg 1947). When the Earth passes through an active meteor stream, radio echoes observed on high frequencies proceed from a narrow zone of the atmosphere of limited depth, which lies in a plane passing through the observing station and perpendicular to the position of the radiant. The rate of occurrence and ranges of the echoes, and the directions of possible reflection points, depend upon the orientation of this plane relative to the aerial coverage, and vary with time as the radiant moves across the sky. The different radiation patterns of the aerial systems of the two equipments result in quite different responses to the same shower.

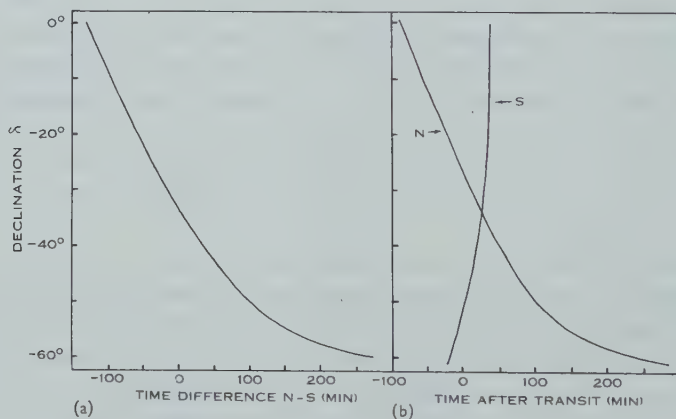


Fig. 1.—Curves for (a) declination, (b) time of transit of a radiant from the times of appearance of echoes of 700 km range in the two aerials (N and S) of the radiant equipment.

For the radiant equipment, the method of radiant determination has been fully described by Clegg (1948) and by Aspinall, Clegg, and Hawkins (1951), and elaboration is unnecessary. By fitting computed range-time envelopes to observed echoes, the time of occurrence of echoes of 700 km range is determined for each aerial. From these times the coordinates of the radiant may be read off immediately from Figure 1, which has been prepared for a mean height of reflection points of 90 km. The N aerial is directed at azimuth 35.5° S. of E.,

the S aerial at 14° N. of E. These curves may be readily extended to north declinations. For radiants more southerly than Dec. -60° echoes do not appear at all in the N aerial, but are detected by the S aerial for many hours and continually if the radiant lies sufficiently close to the south pole of the celestial sphere. Comparison of the theoretical with the experimental range-time envelopes for a given radiant also leads to an estimate of the effective diameter of the radiant area.

A method of radiant measurement, applicable to a wide aperture aerial system, which depends on the variation of the most probable range with time, has been described by McKinley and Millman (1949). This method cannot be applied to the wind equipment because of the small number of suitable echoes. An alternative method which makes use of the known directions of reflection points has therefore been developed for the wind equipment.

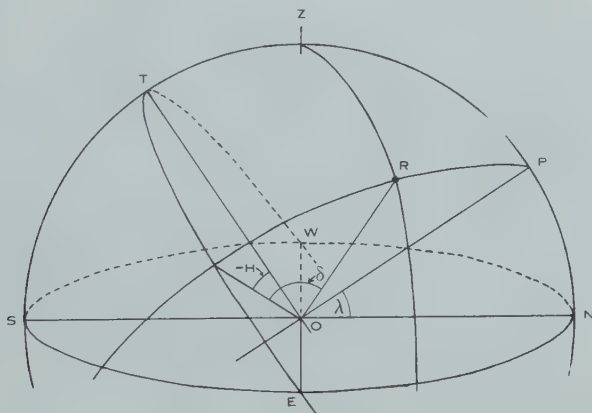


Fig. 2.—Geometry for the determination of radiants by the wind equipment.

In Figure 2, O is the observing station, R the radiant, Z the local zenith, OP the celestial polar axis, and $NESW$ the observer's horizon. Let the direction cosines of the radiant R with respect to the axis system OP , OE , OT be L , M , N ; and those of the reflection point (not shown in the diagram) with respect to these same axes be l , m , n . Then from the condition of specular reflection the equation of the plane containing the observer, the reflection point, and the radiant is

$$lL + mM + nN = 0.$$

If H is the hour angle of the radiant, measured westwards from transit, and δ the declination, then

$$\begin{aligned} L &= \sin \delta, \\ M &= -\cos \delta \sin H, \\ N &= \cos \delta \cos H, \end{aligned}$$

and

$$\tan \delta = \tan \alpha \cos (\varphi + H), \quad \dots \dots \dots (1)$$

where $l = \cos \alpha$, $m/n = \tan \varphi$.

Putting H_0 =hour angle of the radiant at midnight, and t =time of observation in minutes after midnight ($t=0$), then $H=H_0+t/4$ and (1) becomes

$$\tan \delta = \tan \alpha \cos \{H_0 + (\varphi + t/4)\}.$$

This is the polar equation of a straight line in the variables $\tan \alpha = (1-l^2)^{1/2}/l$ and $(\varphi + t/4) = \arctan(m/n) + t/4$, from which the constants $\tan \delta$ and H_0 , which are related to the declination and Right Ascension of the radiant, may be determined.

The direction of the reflection point is actually measured in the axis system NS, EW, OZ . Let these observed direction cosines be l_A, m_A, n_A . The required l, m, n are then found from the observed l_A, m_A, n_A by a simple rotation of axes about EW , thus:

$$\begin{aligned} l &= l_A \cos \lambda + n_A \sin \lambda, \\ m &= m_A, \\ n &= -l_A \sin \lambda + n_A \cos \lambda. \end{aligned}$$

Here λ is the latitude of the observing station.

In the course of radiant determination by this method echoes are grouped into the two categories, shower and sporadic. This grouping is often desirable for such special studies as heights and rates of decay.

Both equipments afford a complete coverage of the visible hemisphere. The accuracy and sensitivity of the radiant equipment are highest for a radiant which transits close to the zenith. On the other hand, an analysis of the echoes indicates that a radiant is most favourably placed for detection by the wind equipment when its elevation is about 30° , so that the sensitivity of this equipment is highest for radiants which transit at this elevation, corresponding to declinations $+25^\circ$ and -85° . The discrimination, as between showers and sporadic echoes, of the wind equipment is inferior to that of the radiant equipment.

With equipments of the wind type, shower activity is readily surveyed by assessing the background (sporadic) rate over periods of shower activity. The shower activity is symmetrical about the time of transit; and, as it has been found that echoes are received from a radiant at very low elevation, the declination δ of the radiant may be found from the relation

$$\cos(T/8) = \tan \lambda \tan \delta,$$

where T is the total time in minutes between onset and cessation of activity and λ the observer's latitude. Radiants determined in this manner do not attain the high accuracy inherent in the two other methods described above but, except in the case of minor showers of short duration, the integration of activity results in a reasonably accurate mean radiant over the duration of the shower.

IV. RADIANTS AND ACTIVITIES OF SHOWERS

(a) *Geminids*

Radio observations (Hawkins and Almond 1952) show that the radiant diameter does not exceed 4° , and the mean radiant and the peak activity remain constant from year to year. This radiant has therefore been used in the calibration of the Adelaide equipments.

Although the Geminid radiant is not favourably situated for observation with the Adelaide radiant equipment, this equipment gave a radiant position for December 14, 1953 of $113 \pm 2^\circ$, $+31 \pm 2^\circ$. The agreement of this radiant with the weighted mean radiant found at the Jodrell Bank Experimental Station for the nights of December 13–15, namely $113.6 \pm 1.2^\circ$, $+31.1 \pm 1.5^\circ$, constitutes an assurance that the error in radiant coordinates due to aerial orientation at Adelaide will not exceed 1° . Uncorrected hourly echo rates, based on the echoes falling within the range-time envelopes, are shown in Figure 3 (a).

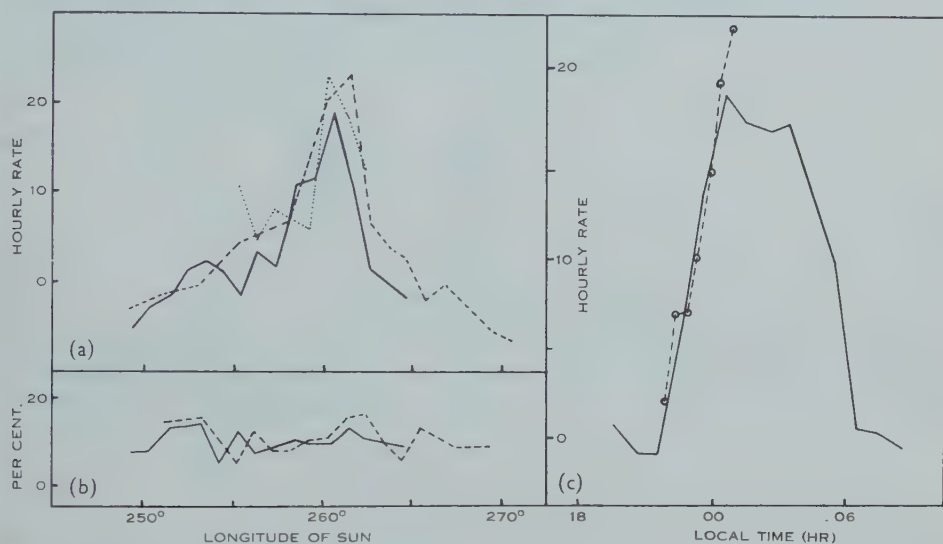


Fig. 3.—Echo rates for the Geminid shower. (a) Rate of shower echoes for each day, averaged over the period 22–06 hr, wind equipment, — 1952, --- 1953. Rate of echoes, for each day, falling within range-time envelopes, radiant equipment, ... 1953. (b) Percentage of all echoes (background+shower) whose duration exceeds 2 sec. (c) — Hourly rate, for each hour, of shower echoes averaged over December 11–14, 1952 and 1953, wind equipment. ○ --- ○ visual half-hourly rates, 1952.

The shower echo rate as a function of the time of day (00 hr=midnight) for December 11–14 inclusive, as recorded by the wind equipment, is illustrated in Figure 3 (c), in which data for 1952 and 1953 are combined. These rates lead to a radiant position 110° , $+30^\circ$. Day-to-day radiant coordinates found in 1952 by the more rigorous method are as follows:

December 11	111° , $+27^\circ$
„ 12	111° , $+26^\circ$
„ 13	113° , $+29^\circ$
„ 14	115° , $+29^\circ$

The activity observed on each day, which is the shower activity averaged over the 8 hr from 2200 to 0600 inclusive, is shown in Figure 3 (a). The peak of the activity falls between $\odot=260^\circ$ and $\odot=261.5^\circ$, which is to be compared with other determinations by Hawkins, 261.1° , and Hoffmeister, 261.5° (Hawkins and Almond 1952).

Figure 3 (b) indicates the percentage of echoes whose duration exceeds 2 sec. This information is relevant to the study of the structure of the Geminid stream and is considered in this context in Section V.

TABLE 1
RADIANTS OF DAY-TIME SHOWERS 1953

Date	Arietids		ζ -Perseids	
	R.A.	Dec.	R.A.	Dec.
June 5 ..	46°	+21°	65°	+21.5°
6 ..	Low activity		62°	+23°
7 ..	Not resolvable		Not resolvable	
8 ..	42°	+22°	68°	+23°
9 ..	44°	+23.5°	66°	+22.5°
10 ..	45°	+24°	69°	+21°
11 ..	48°	+19°	Low activity	
12 ..	47°	+22°	66°	+26°

(b) June Day-time Showers

These showers have been observed during 1952 and 1953 with the wind equipment. The hourly echo rates are given in Figure 4 (c); after noon gaps appear in the records for some days and the descending portions of the curves

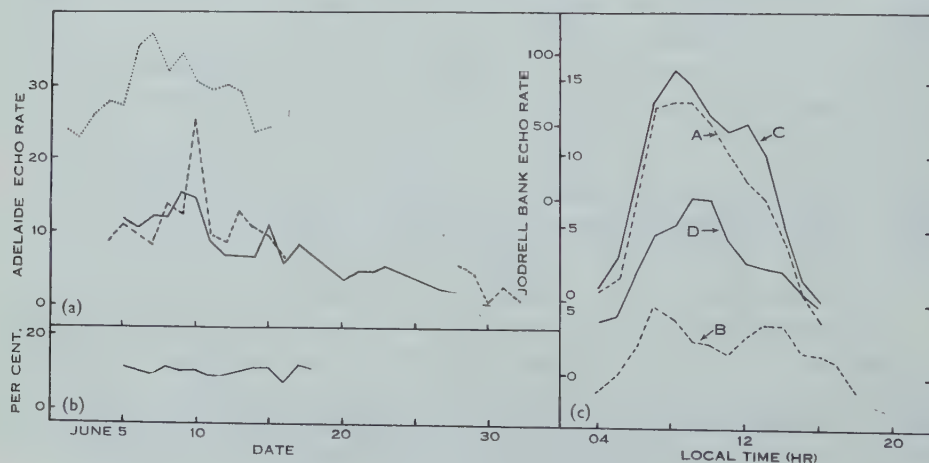


Fig. 4.—Echo rates for the June day-time showers. (a) Rate of shower echoes for each day averaged over the period 05-16 hr, wind equipment, ---- 1952, — 1953. Rate of echoes, for each day, falling within range-time envelopes, Jodrell Bank radiant equipment, 1952. (b) Percentage of all echoes whose duration exceeds 2 sec, 1953, wind equipment. (c) Hourly rate, for each hour, of shower echoes averaged over the following periods: curve A, June 1-16, 1952; curve B, June 28-July 3, 1952; curve C, June 4-18, 1953; curve D, June 20-28, 1953.

are not considered to be as reliable as the ascending portions. The shape of the rates curves may also be influenced by the diffuse background (Lovell and Clegg 1952, p. 95). The Arietid and ζ -Perseid radiants cannot be resolved by

this analysis. The activity from June 28 to July 3, 1952 proceeds from the β -Taurid radiant. Day-to-day radiant positions for the two main showers during 1953 are given in Table 1.

The day-to-day activity on the wind equipment, averaged over a period of 11 hr, is shown in Figure 4 (a). For comparison, the sum of the activities of the Arietids and ζ -Perseids for 1952, as measured at Jodrell Bank by Almond, Bullough, and Hawkins (1952) over the same days, using an equipment of the radiant type, is reproduced. Further investigation is necessary before ascribing the very decided differences between these two measures of the activity to real (and large) short-period irregularities in the activities of these streams, rather than to different instrumental techniques.

TABLE 2
RADIANTS OF δ -AQUARID SHOWER 1953

Date	R.A.	Dec.	Hourly Rate	Type of Radiant*	No. of Echoes > 600 km
July 21 ..	336°	-17°	18	D	0
22 ..	336°	-17°	19	C	3
23 ..	337°	-17°	—		—
24 ..	338°	-17°	33	D	3
25 ..	338°	-17°	41	C	4
26 ..	339°	-17°	52	D; Sc; Second centre	3
27 ..	340°	-17°	54	C; Sc; <3°	5
28 ..	340°	-18°	41	D; Sc; <3°	0
29 ..	340°	-18°	48	C; Sc; <3°	2
30 ..	342°	-16°	32	D	3
31 ..	343°	-16°	44	D; Sc;	6
Aug. 1 ..	344°	-16°	43	C; Sc; <3°	0
2 ..	345°	-15°	50	VC; 3°	8
3 ..	346°	-15°	26	D	6
4 ..	347°	-14°	21	C	0
5 ..	348°	-14°	33	VD	8
6 ..	348°	-14°	21	D	4
7 ..	348°	-15°	18	D	2
8 ..	347°	-17°	12	C	1

* C=compact, a well-defined radiant with little activity outside the envelope; D=diffuse, considerable activity on either side of the main radiant; Sc=strong concentration, a concentration of echoes to the theoretical envelope which implies a rather small radiant area.

(c) δ -Aquarids

The visual workers have established that the period from the beginning of July to the middle of August is one of high meteor activity in the southern hemisphere. The principal shower is the δ -Aquarids, but McIntosh (1935) lists no less than 43 minor radiants in the range of Right Ascension from 300° to 11°, and of declination from -8° to -33°, during this period. The most active of these minor streams are the ζ -Aquarids, δ -Capricornids, and α -Pisces Australids. In view of the low resolution of the radio equipments, as opposed to visual

methods, it is not to be expected that these minor streams will be resolvable by the radiant equipment, and this has in fact been confirmed.

The position, activity, and nature of the δ -Aquirid radiant for 1953, as measured with the radiant equipment, are listed in Table 2. The hourly echo rate is an average over the number of echoes falling within fitted range-time envelopes. The column headed "Type of Radiant" lists the nature of the activity, thus: C—compact, a well-defined radiant with little activity outside the envelope; D—diffuse, considerable activity on either side of the main radiant; Sc—strong concentration, a concentration of echoes to the theoretical envelope which implies a rather small radiant area. This method of dealing with the problem of the size of the radiant is forced by the presence of the diffuse

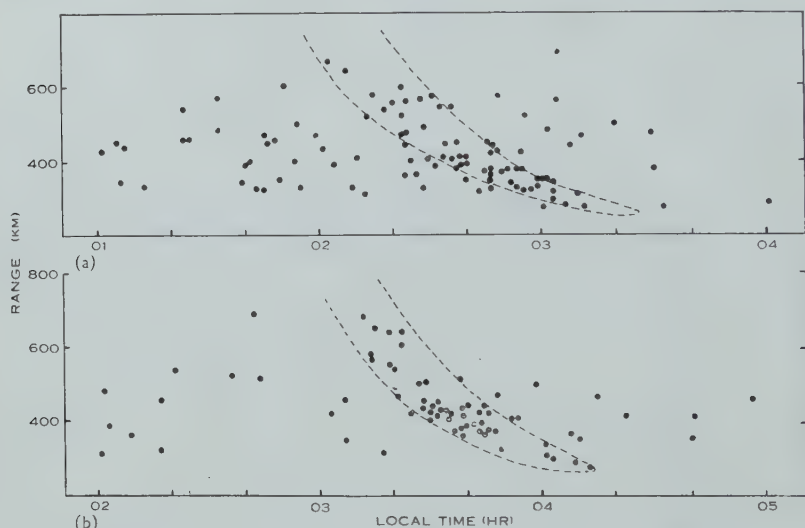


Fig. 5.—Observed echoes (·) and fitted theoretical range-time envelopes (---) for radiant equipment, δ -Aquirids 1953. (a) July 26, N aerial; (b) July 27, S aerial.

activity, but when the diameter of the radiant area could be determined it appears in Table 2. Finally, the number of echoes whose range exceeds 600 km is an index of the reliability of the radiant, a large number of long-range echoes implying an accurate radiant. Two experimental range-time plots with fitted envelopes are drawn in Figure 5 to illustrate the nature of the classification problem.

These radiant positions agree closely with earlier visual determinations by McIntosh (1935) and Hoffmeister (1948). Radar determinations in the northern hemisphere (Hawkins and Almond 1952; Lindblad 1952), which give too northerly declinations, are apparently influenced by the diffuseness of the radiant and its low elevation.

At the edges of the δ -Aquirid stream, i.e. for several days prior to July 21 and after August 8, the activity was confused. It was spread more or less uniformly over a period of several hours each night and the radiant was diffuse

and ill-defined. Whilst the radiant had contracted by July 21, over the next few days subsidiary active centres appeared, with a possible radiant at 342° , -18° active on July 21–22. Over the peak of the activity, from July 27 to August 2, the δ -Aquarid radiant proper was compact, its diameter usually being less than 3° . The diffuse activity attributed to the minor streams appears intermittently throughout the duration of the main shower. Its importance is well brought out in Figure 6, which compares the hourly rate of activity attributed to the δ -Aquarid radiant proper with the total number of echoes detected each day over a period of 3 hr centred on the times of peak activity in the respective aerals.

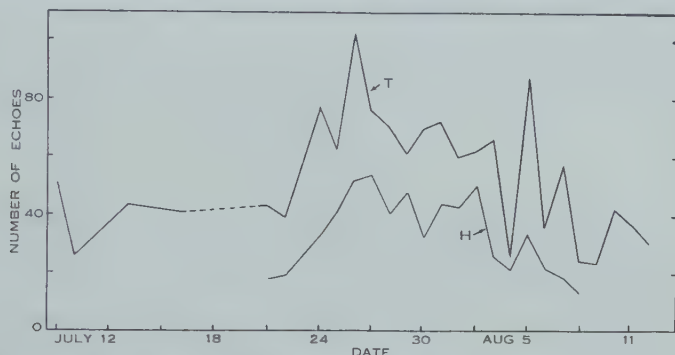


Fig. 6.—Echo rates for δ -Aquarids 1953, radiant equipment. Curve *T*, total number of echoes detected each day in the 3-hr intervals from 01 to 04 hr N aerial, and from 02 to 05 hr S aerial; curve *H*, hourly rate of δ -Aquarid echoes for each day.

(d) *Corona Australids*

Figure 7 depicts the echo rates for this shower as recorded by the wind equipment in 1953.

The radiant position, estimated as 250° , -50° , cannot be fixed precisely because of the low activity and also because of the marked deficiency of large meteors in this stream (Section V). As the radiant culminates after sunrise, very few visual observations are available. McIntosh (1935) lists five radiants falling between 272° , -40° and 283° , -38° .

Clearly an accurate radar fix on this radiant is essential.

(e) *Orionids*

The Orionids 1952 appeared on the wind equipment with peak activity 20/hr on October 25. No shower was detected during 1953. Similar variability has been reported by Hawkins and Almond (1952).

(f) *Other Showers*

Two minor showers have been detected at the limit of resolution of the radiant equipment. The first, giving an echo rate of 7/hr, was active over August 26–29, 1953, at a radiant position 03° , -14° ; the second over September 6–7, 1953, at 02° , -53° with an activity of 11/hr.

At the end of May and the beginning of July 1953 the radiant equipment operated intermittently on one aerial only. Two radiants were active over the

period May 26–31, giving echoes in the S arial at 04 and 12 hr respectively, but unfortunately the range-time plots were too confused for even approximate radiant positions to be deduced with the single arial. No visual activity during this period has been listed.

Day-time activity was observed on the S arial on July 9, 10, and 13 between 07 and 10 hr, and again on the N arial on July 16 from 08 to 11 hr, the rate being about 10/hr. However, the gaps in the records are too wide to justify the assumption that this activity proceeds from only one radiant.

As already mentioned, no major day-time showers active in the southern hemisphere have yet been detected.

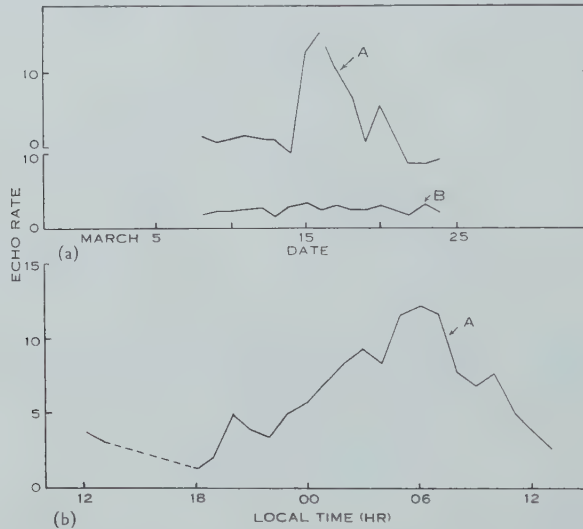


Fig. 7.—Echo rates for Corona Australids 1953, wind equipment. The rate of shower echoes for each day is an average over the period from 00 to 11 hr. Curve A, all shower echoes; curve B, echoes used in wind measurement. The hourly rate, for each hour, of shower echoes is an average over March 15–20, 1953.

V. THE MASS DISTRIBUTIONS OF METEOR STREAMS

(a) *Rate of Detection of Meteors*

After making due allowance for the striking variability in the activity of a given shower from day to day and from year to year, the activities observed at different latitudes, or using different techniques, are not the same. Some interesting information regarding the structure of meteor streams emerges from an attempt to reconcile these various rates.

The visual activity of a shower depends of course upon the incident flux of the meteors and their mass distribution. In addition, the observed visual rate is a function of the zenith angle z of the radiant, the relation being

$$N_v = 1.007 N_0 \cos (z - 6.5^\circ), \quad \dots \dots \dots (2)$$

where N_0 is the zenithal rate (Olivier 1925, p. 195).

The radio echo rates are influenced by factors other than the mass distribution and the ionizing efficiency. These are (i) the collecting area of the aerial system, which may vary widely with radiant position; and (ii) the limiting sensitivity of the equipment, which usually corresponds to meteors lying below the visual limit.

Starting with the mass distribution

$$v_m dm = \text{const. } m^{-p} dm, \quad \dots\dots\dots (3)$$

Kaiser (1953) has calculated the observed echo rate in terms of the properties of the incident meteor flux and the radio equipment parameters. Applying his formula to the radiant equipments at Adelaide and at Jodrell Bank, whose main aerial lobes are almost axially symmetrical, the dependence of the echo rate upon the zenith angle is found to be approximately

$$N_R = F(p) \cos^{p-1} z \quad \dots\dots\dots (4)$$

for radiants for which $z < 70^\circ$. The function $F(p)$ depends upon the equipment parameters, which are constant for a given equipment.

The expression (4) is not applicable to the wide aperture aerial system of the wind equipment. For low radiant elevations shower meteors are collected near the zenith, where aerial gain and collecting area are relatively insensitive to zenith angle, and the shower rate for the wind equipment will follow a law of the form

$$N_W = \text{const. } \cos^n z, \quad \dots\dots\dots (5)$$

where the exponent n is close to unity. Actual values found for n range from 0.8 to 1.1. The close correspondence between the relations (2) and (5) is well brought out by some visual observations at Adelaide on the Geminid shower 1952. These observations, made on December 13 and 15, are shown in Figure 3 (c). Individual points represent half-hourly rates for the whole of the visible hemisphere, watched by four observers.

Activities of the major showers observed visually in the northern hemisphere, by the radiant equipment at Jodrell Bank, and by the two Adelaide equipments, have been freed from the influence of zenith angle by use of (2), (4), and (5). The values of p found by Kaiser (1953) were adopted. Any differences remaining between the four sets of activities for individual showers must then be the result of different incident fluxes, which in turn are determined by the mass distribution (3) as well as by the limiting sensitivities of the equipments.

Consistent zenithal rates were found for the Geminids ($p=1.7$) and the Perseids ($p=1.6$). The Adelaide wind rates were much less than the Jodrell Bank radiant rates for the June day-time showers, which appears to confirm the deficiency of meteor particles of larger masses implied by the value of $p=2.25$ for the Arietid stream. The radiant equipment rates for the δ -Aquadrid shower exceed the visual rate by a factor of 3, which again may imply a large value of p . This is suggestive in view of the proposal by Almond (1952) that the δ -Aquadrid and the day-time Arietid showers are produced by one extended meteor stream.

(b) Durations of Echoes

According to Kaiser (1953), the distribution of durations of echoes of the long-enduring type, observed with a sufficiently sensitive equipment, is

$$v_t \propto t^{-(3p+1)/4}, \quad \dots\dots\dots (6)$$

where t is the duration of an individual echo, and v_t is the number of echoes per duration increment δt . The mass distribution may thus be found from an analysis of the durations of individual echoes, and the wind equipment records have been examined to this end. Distributions have been determined for the Geminids 1952 and 1953, and the June day-time showers 1953. Assuming that the observed distributions may be represented by the power law

$$v_t \propto t^{-s},$$

the values of s are :

Geminids	..	Shower 1952	2.14
		Shower 1953	1.44
		Background	2.09
June showers	..	Shower 1953	2.54
		Background	1.73

For sporadic meteors $p=2.0$ and $(3p+1)/4=1.75$. The agreement between this value and the values of s given above for background echoes is satisfactory, in view of the undoubted admixture, into the echoes used in deriving these values of s , of a number of short, decay type, echoes.

The high value of s for the June day-time showers is consistent with the deficiency of large meteors already inferred. Figure 4 (b), in which the percentage of all echoes whose duration exceeds 2 sec is plotted over the duration of the two principal showers, indicates that the structure of these streams is uniform in cross section.

The value of $s=1.44$ for the Geminids 1953, and Figure 3 (b), indicate the existence of an asymmetrical core of heavy meteors within this stream, a conclusion already reached by Hawkins and Almond (1952). However, there is no evidence for such a core in 1952.

Finally, Figure 7 (a) shows the day-to-day activity over the duration of the Corona Australids, as assessed by the numbers of echoes actually used in wind measurement. Although selection of echoes is involved, this sample is thought to afford a reliable indication of the incidence of echoes whose duration exceeds $\frac{1}{2}$ sec. Although at the peak of the activity the total number of shower echoes detected exceeds the background rate, there are evidently very few shower echoes of the long-enduring type. This is confirmed by calculation of the hourly rates for the same group of echoes used in wind measurement; the total hourly rate averaged over shower days differs from that for the non-shower days by 0.1 ± 0.8 . The Corona Australid stream is apparently quite deficient in large meteors.

VI. THE BACKGROUND ECHOES

The background echoes represent the sum of true sporadic meteors and the minor showers which are not resolvable by the radio equipments. The following discussion is based on over 36,000 echoes.

Typical diurnal rates of background echoes are illustrated in Figure 8, for the month of December 1953. Figure 10 shows the seasonal variation of the mean monthly hourly rates, normalized to December. The higher background rate of the wind equipment, despite its lower sensitivity, is due to the much larger collecting area of its aerial system.

Detailed examination of the available material suggests that the seasonal change in the background rate for the wind equipment is simply an alteration in the scale of the diurnal rate. Neither the ratio of maximum to minimum diurnal rate, nor the times of these maxima and minima, depend significantly upon the season. For the radiant equipment, the extent of the diurnal variation and the mean monthly hourly rate appear to be almost independent of season. The diurnal variation in the background rate for the radiant equipment is

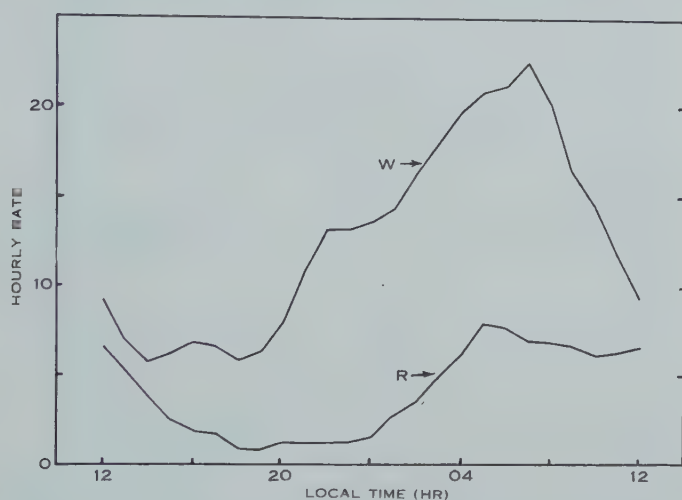


Fig. 8.—Diurnal variation in background rates, December 1953.

Curve R, radiant equipment; curve W, wind equipment.

almost twice that for the wind equipment, the respective ratios of maximum to minimum being 8.3 and 4.4. It should be remembered that the radiant equipment observations are confined to the latter half of the year. The shapes of the diurnal curves for the two equipments are dissimilar; for the wind equipment there is a tendency to sharp maxima near 06 hr and flat minima near 18 hr, whilst for the radiant equipment the maxima are flat and the minima much sharper. The explanation of these facts lies in the different radiation patterns of the two aerial systems.

The geocentric velocity and the apparent radiant of any meteor are the result of compounding the true (heliocentric) velocity of the meteor with the velocity of the Earth, which is directed towards the apex of the Earth's way. If a uniform distribution of radiants and activity of background meteors is assumed and the dependence of luminous and ionizing efficiency on zenith angle is ignored, the seasonal and diurnal variations in the background rate may be evaluated without difficulty. The method of calculation of these rates for

the visible hemisphere has been given in detail by Davidson (1914). These calculations are readily extended to meet the case of the more limited collecting areas of the radio equipments.

An analysis of the directions of the reflection points for echoes detected by the wind equipment shows that the very great majority of the echoes proceeds from zenith angles between 20° and 40° and that there is no seasonal variation. For this equipment, therefore, the apparent radiants may be taken to lie in a

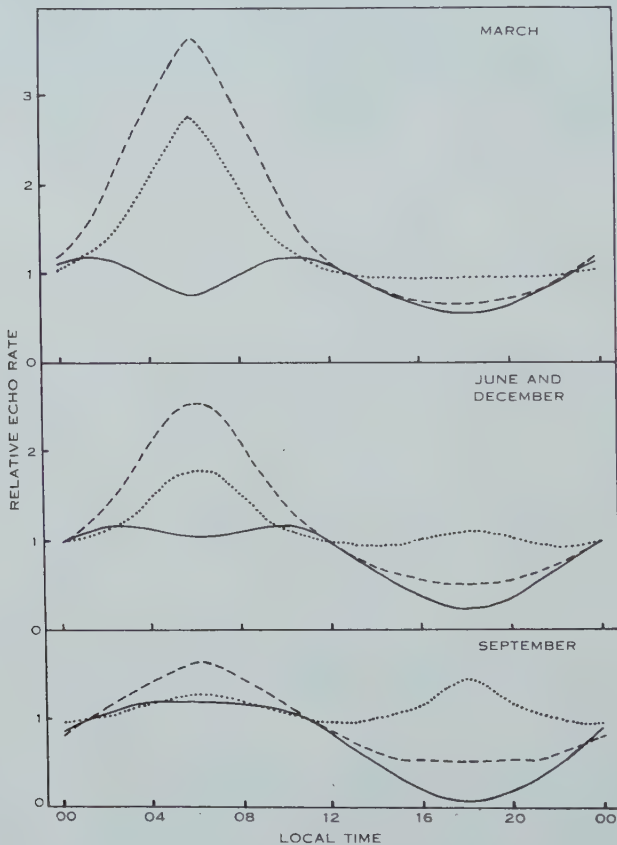


Fig. 9.—Diurnal variation in the rate of background echoes at the equinoxes and the solstices, for various collecting areas. ——— 30° spherical cap; ---- strip comprising elevations 20° – 40° ; strip comprising elevations 0° – 20° .

narrow strip of the visible hemisphere lying between elevations 20° and 40° . For the radiant equipment the majority of apparent radiants are located near the zenith, and a spherical cap of radius 30° is a reasonable first approximation to the zone of apparent radiants. The diurnal variations in activity associated with these two areas of the sky are shown in Figure 9 for the latitude of Adelaide, at the equinoxes and the solstices, on the assumption of parabolic velocities for meteors.

It will be seen that the characteristic shapes of the observed diurnal variations for the two equipments are reproduced by these two geometries, as is also the greater diurnal range for the radiant equipment over the latter half of the year. The small seasonal variation in the ratio of maximum to minimum diurnal rate is also predicted by the model for the wind equipment, but in view of the assumptions made in the derivation of the theoretical curves it would not be wise to press such comparisons.

Seasonal variations in the mean monthly hourly rate expected on the two geometries, along with the observed values, are shown in Figure 10. Incidental calculations suggest that a strong seasonal variation is not to be expected with a collecting area of any shape. The increase in activity observed from June

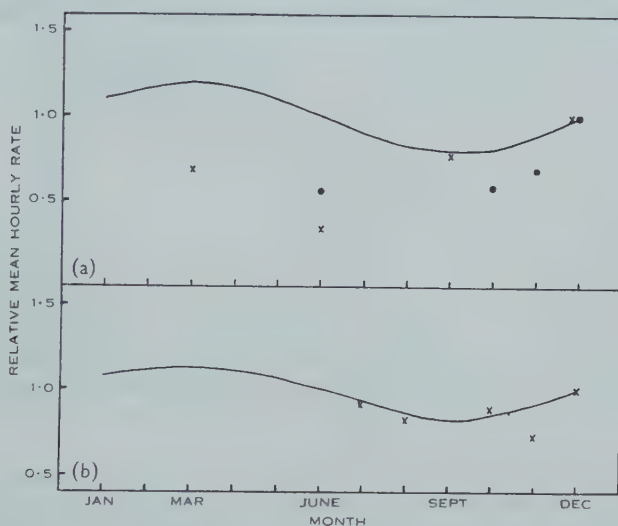


Fig. 10.—Seasonal variation in the relative mean monthly hourly rate of background echoes. (a) Wind equipment 1952 (·) and 1953 (×); (b) Radiant equipment 1953 (×). Expected variations are shown by full lines.

to December, for the wind equipment, occurred in both 1952 and 1953, and appears to be due to a real increase in background activity towards the end of the year. However, this is not confirmed by the limited observations with the radiant equipment, nor by visual observations of McIntosh (Hoffmeister 1937, p. 60) in the southern hemisphere, although according to Olivier (1925) the real visual background activity in the northern hemisphere increases towards the end of the year. The visual observations contain ample evidence of discrepancies from observer to observer, and further sampling by radio techniques is necessary.

The diurnal variation of activity expected from apparent radiants lying between elevations of 0 and 20° is shown in Figure 9. Although no allowance has been made here for the rapid falling off of ionizing efficiency at very low elevations, it appears that a valuable contribution to this problem may be made by special radio equipments designed for the purpose of counting sporadic, rather than shower, meteors.

VII. HEIGHTS OF REFLECTION POINTS

The heights of reflection points of over 3000 echoes detected with the wind equipment have been computed from measured directions and slant ranges. Individual heights are accurate to ± 2 km.

(a) *Showers*

The height distribution of approximately 100 Geminid echoes is shown in Figure 11 (c). This distribution was determined by a statistical method; but a smaller sample of echoes, known by the sorting method described in Section III to belong to Geminid meteors, gave an identical distribution.

The height distributions of Figures 11 (a) and (b) refer to the ζ -Perseids and the day-time Arietids. The most probable height of 89 km for the Arietids is much higher than the height of 80 km found by Clegg and Davidson (1950), but less than the more recent value of 92.5 km derived from decay times (Kaiser

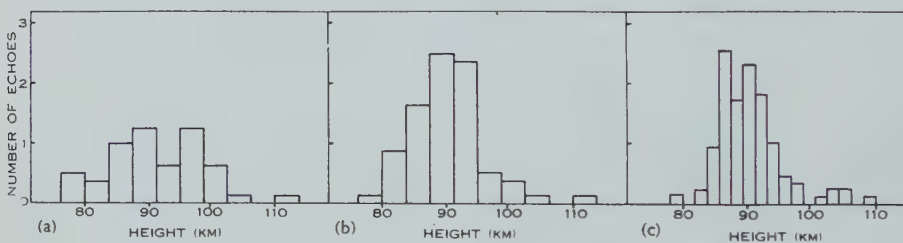


Fig. 11.—Height distributions of shower echoes. (a) ζ -Perseids 1953 (47 echoes); (b) day-time Arietids 1953 (69 echoes); (c) Geminids 1952 (100 echoes).

1953). The height of maximum ionization of a meteor trail depends on the zenith angle χ of the meteor path, the relation being approximately $h - h_0 = 6 \ln \cos^{2/3} \chi$. The height change corresponding to a change of zenith angle from e.g. 60 to 30° is 2 km. This will be reflected in a dependence of the height of the reflection point upon χ , but as the zenith angles for the Jodrell Bank echoes are not known the correction cannot be assessed.

(b) *Sporadic Meteors*

Background echoes observed during March, June, September, and December 1953 have been analysed in 6-hourly groups centred on 00, 06, 12, and 18 hr.

The diurnal variations in most probable heights are shown in Figure 12. Furthermore, the height distribution is much sharper at 12 hr than at 00 hr, whilst at 06 and 18 hr the distribution is intermediate. No seasonal effects were found, either in most probable heights or in the shapes of the distributions. The general nature of the height distributions is sufficiently illustrated by those for some other months already given by Elford and Robertson (1953).

Meteor heights are strongly influenced by the geocentric velocity, which depends on the elongation of the apparent radiant, i.e. the angular distance of the apparent radiant from the apex of the Earth's way. In March at Adelaide the apex reaches an elevation of 79°, whilst in September the maximum elevation is only 33°. Height distributions taken near 06 hr in March and 18 hr in

September would be expected to be sharper than those taken at 18 hr in March and 06 hr in September. On this simple picture, also, maximum heights are expected at 06 hr and minimum heights at 18 hr, whilst the distributions at 00 and 12 hr should be identical. These expectations are not confirmed by the observed distributions.

The selection of echoes used in this analysis is partly conditioned by the wind pattern in the meteor zone, but no relation between wind velocity and measured heights (other than the wind gradient) could be found. There are no marked diurnal or seasonal changes in the zenith angles of reflection points.

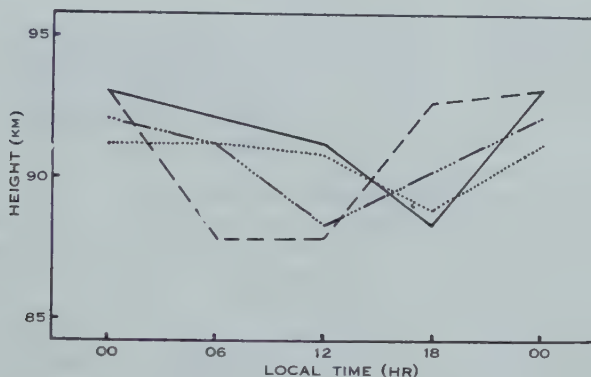


Fig. 12.—Diurnal variation of the most probable height of background echoes. — March; - - - June; - · - · - September; · · · · · December.

In any case the occurrence of the maximum diurnal echo rate at 06 hr and of the minimum at 18 hr renders any explanation in terms of zenith angles improbable. Nevertheless the distinctive features of the height distribution are thought to be real and probably astronomical in origin, although the possibility of atmospheric influence cannot be ruled out.

VIII. ACKNOWLEDGMENTS

The author wishes to thank Professor L. G. H. Huxley for his constant encouragement and advice, and Professor A. C. B. Lovell of the Jodrell Bank Experimental Station for his joint sponsorship of the project. He is particularly indebted to Mr. W. G. Elford for the benefit of many valuable discussions on the interpretation of the data, for the development of the method of radiant determination for the wind equipment, and for reduction of observations. The height data were analysed by Mr. J. W. Smith. Members of the Astronomical Society of South Australia collaborated in the visual observations in December 1952.

Mr. R. J. Anglely assisted in the construction of the radiant equipment, part of which was kindly supplied by the University of Manchester. The author also wishes to thank the Department of Supply for providing the field site at Salisbury and for assistance with equipment.

This work was financed by the University of Adelaide and the Radio Research Board, C.S.I.R.O., to whom the author is also indebted for maintenance grants.

IX. REFERENCES

- ALMOND, M. (1952).—*Jodrell Bank Ann.* **1** : 22.
 ALMOND, M., BULLOUGH, K., and HAWKINS, G. S. (1952).—*Jodrell Bank Ann.* **1** : 13.
 ASPINALL, A., CLEGG, J. A., and HAWKINS, G. S. (1951).—*Phil. Mag.* **42** : 504.
 CLEGG, J. A. (1948).—*Phil. Mag.* **39** : 577.
 CLEGG, J. A., and DAVIDSON, I. A. (1950).—*Phil. Mag.* **41** : 77.
 DAVIDSON, M. (1914).—*J. Brit. Astr. Ass.* **24** : 352.
 DAVIES, J. G., and ELLYETT, C. D. (1949).—*Phil. Mag.* **40** : 614.
 ELFord, W. G., and ROBERTSON, D. S. (1953).—*J. Atmos. Terr. Phys.* **4** : 271.
 HAWKINS, G. S., and ALMOND, M. (1952).—*Mon. Not. R. Astr. Soc.* **112** : 219.
 HEY, J. S., and STEWART, G. S. (1947).—*Proc. Phys. Soc. Lond.* **59** : 858.
 HOFFMEISTER, C. (1937).—"Die Meteore." (Akademische Verlagsgesellschaft : Leipzig.)
 HOFFMEISTER, C. (1948).—"Meteorströme." (Weimar.)
 KAISER, T. R. (1953).—*Advanc. Phys.* **2** : 495.
 LINDBLAD, B. (1952).—Chalmers Tek. Högsk. Handl. No. 129.
 LOVELL, A. C. B., BANWELL, C. J., and CLEGG, J. A. (1947).—*Mon. Not. R. Astr. Soc.* **107** : 164.
 LOVELL, A. C. B., and CLEGG, J. A. (1948).—*Proc. Phys. Soc. Lond.* **60** : 491.
 LOVELL, B., and CLEGG, J. A. (1952).—"Radio Astronomy." (Chapman and Hall : London.)
 MCINTOSH, R. A. (1935).—*Mon. Not. R. Astr. Soc.* **95** : 709.
 MCKINLEY, D. W. R., and MILLMAN, P. (1949).—*Proc. Inst. Radio Engrs., N.Y.* **37** : 364.
 OLIVIER, C. P. (1925).—"Meteors." (Williams and Wilkins : Baltimore.)
 ROBERTSON, D. S., LIDDY, D. T., and ELFord, W. G. (1953).—*J. Atmos. Terr. Phys.* **4** : 255.

THE GRAVITATIONAL COMPRESSION OF AN ELASTIC SPHERE

By A. KEANE*

[Manuscript received August 19, 1954]

Summary

On considering a sphere in hydrostatic gravitational equilibrium, composed of a homogeneous elastic material for which the variation of incompressibility κ with pressure p is given by $d\kappa/dp = n$, a constant, we find that there is an upper bound to the radius R of the sphere provided $n \leq 2$, and that for all values of n there is a lower bound to the value of I/MR^2 , where I is the moment of inertia about a diameter and M is the mass of the sphere.

For a sphere composed of material for which $d\kappa/dp$ is a decreasing function of pressure and $d\kappa/dp \rightarrow 5/3$ as $p \rightarrow \infty$, it emerges that $0.40 > I/MR^2 > 0.23$ and that the maximum radius is of the order of 10^4 km.

I. INTRODUCTION

Seventy years ago the major planets were either regarded as "water planets" or else they were thought to consist of a core of terrestrial material surrounded by a gaseous atmosphere. The acceptance of these hypotheses was due to the lack of knowledge on the compressibility of solid and liquid materials, and, although both hypotheses are now known to be incorrect, the effect of compression is still not completely known.

The theories of quantum mechanics and finite elastic strain, together with the high pressure experiments of Bridgman, have yielded results which should ultimately lead to the exact solution of the density distribution throughout a self-compressed elastic sphere. The analytical solution of the problem is forestalled by the lack of an exact solution of the Lane-Emden equation, but certain indicative results will be obtained in this paper by making use of the few cases in which the Lane-Emden equation has an analytical solution.

Using the results of quantum mechanical calculations and solving the Lane-Emden equation by numerical integration, Ramsey (1950) has solved the problem of the density distribution throughout a hydrogen planet. Earlier Birch (1939) used finite elasticity theory combined with numerical integration to determine the density distribution throughout a homogeneous layer of the Earth.

The model with which this paper is concerned is a homogeneous sphere composed of material for which $d\kappa/dp$ is constant, and emphasis is placed on determining the lower bound of the ratio I/MR^2 and the upper bound of the radius. In the discussion at the end of this paper it is pointed out that the results can be applied to a sphere of material for which $d\kappa/dp$ is a monotonic decreasing function of pressure.

* Wollongong Technical College, Wollongong, N.S.W.

II. THE DIFFERENTIAL EQUATION FOR DENSITY

Using the equation defining incompressibility, namely, $\kappa = \rho(dp/d\rho)$, where ρ is the density at pressure p , together with the relation $\kappa = \kappa_0 + np$, where κ_0 is the incompressibility at zero pressure, we obtain the pressure-density relation

$$p = \frac{\kappa_0}{n} \left\{ \left(\frac{\rho}{\rho_0} \right)^n - 1 \right\}, \quad \dots\dots\dots (1)$$

where we have used ρ_0 to denote the density at zero pressure.

Combining (1) with the equation for hydrostatic equilibrium of a sphere, namely,

$$\frac{1}{r^2} \frac{d}{dr} \left(\frac{r^2}{\rho} \frac{dp}{dr} \right) = -4\pi G \rho, \quad \dots\dots\dots (2)$$

where G denotes the gravitation constant and r is distance from the centre of the sphere, we obtain the differential equation

$$\frac{1}{r^2} \frac{d}{dr} \left\{ r^2 \frac{d}{dr} \rho^{n-1} \right\} = - \frac{4\pi G \rho_0^n (n-1)}{\kappa_0} \cdot \rho. \quad \dots\dots (3)$$

III. THE DENSITY DISTRIBUTION FOR $n=2$

On denoting $4\pi G \rho_0^2 / \kappa_0$ by β^2 and making the substitution $\rho = \chi/r$ we find that the general solution of (3) is $\rho = C \sin(\beta r - \delta)/r$ where C and δ are arbitrary constants of integration. Since density is to be finite at the origin we have $\delta=0$, and, on considering the surface values of a sphere of radius R ,

$$\frac{\rho}{\rho_0} = \frac{R}{r} \frac{\sin \beta r}{\sin \beta R}. \quad \dots\dots\dots (4)$$

Since the mass of the sphere $M = \int_0^R 4\pi r^2 \rho(r) dr$ we have on substituting from equation (4) and integrating, that

$$\frac{1}{3} \frac{\bar{\rho}}{\rho_0} = \frac{1}{\beta^2 R^2} - \frac{\cot \beta R}{\beta R}, \quad \dots\dots\dots (5)$$

where $\bar{\rho}$ is the mean density given by $M = 4\pi R^3 \bar{\rho}/3$. Similarly, from the formula

$$I = \frac{8\pi}{3} \int_0^R r^4 \rho(r) dr,$$

where I is the moment of inertia of the sphere about a diameter, combined with equation (4), we deduce that

$$\frac{1}{2} a \frac{\bar{\rho}}{\rho_0} = - \frac{\cot \beta R}{\beta R} + \frac{3}{\beta^2 R^2} + \frac{6 \cot \beta R}{\beta^3 R^3} - \frac{6}{\beta^4 R^4},$$

where a is given by $I = aMR^2$, and on substituting in this equation from (5) we obtain

$$a = \frac{2}{3} + \frac{4}{3(1 - \beta R \cot \beta R)} - \frac{4}{\beta^2 R^2}. \quad \dots\dots\dots (6)$$

IV. APPROXIMATE SOLUTION OF EQUATION (3) FOR ARBITRARY $dn/d\rho$

As a first approximation to the density throughout the sphere we take $\rho = \bar{\rho}$, where $\bar{\rho}$ is the mean density of the self-compressed sphere. On solving equation (3) and remembering that ρ must be finite at the origin, we obtain as a second approximation to the density

$$\rho = \rho_0(A - Br^2)^\alpha, \quad \dots\dots\dots (7)$$

where $A = 1 + BR^2$, $B = \frac{2}{3} \pi G \rho_0 \bar{\rho} / \kappa_0 \alpha$, and $\alpha = 1/(n-1)$.

On determining the mass and moment of inertia as in the previous section, we obtain the results

$$\frac{\bar{\rho}}{\rho_0} = A^\alpha F\left\{-\alpha, \frac{3}{2}; \frac{5}{2}; \frac{BR^2}{A}\right\}, \quad \dots\dots\dots (8)$$

$$\alpha \frac{\bar{\rho}}{\rho_0} = \frac{2}{5} A^\alpha F\left\{-\alpha, \frac{5}{2}; \frac{7}{2}; \frac{BR^2}{A}\right\}, \quad \dots\dots\dots (9)$$

where $F\{a, b; c; z\}$ is the hypergeometric function.

It is necessary at this stage to give some idea of the range of validity of the equations (7), (8), and (9).

On substituting the approximate solution (7) into the differential equation (3) we find that the left side is constant while the right side varies by the fraction $(A - BR^2/A)^\alpha$ of itself as r varies from 0 to R . Since $(A - BR^2/A)^\alpha = 1/A^\alpha = \rho_0/\rho_c$, it follows that the solution (7) will be a valid approximation only if the density variation throughout the sphere is not large, so that ρ_0/ρ_c is approximately unity. We use ρ_c to denote central density.

From this argument it appears to follow that equations (8) and (9) are not valid for large compression because of their dependence on (7). This, however, is not the case since equations (8) and (9) are obtained from (7) by a further use of the equations of hydrostatic equilibrium and are therefore true to a higher order of approximation.

Let us consider the derivation of (8) as follows.

The equations of hydrostatic equilibrium are

$$\frac{dp}{dr} = -\frac{GM(r)}{r^2} \rho, \quad \frac{dM(r)}{dr} = 4\pi r^2 \rho,$$

where $M(r)$ is the mass within the sphere of radius r , and on using (1) we can derive the differential equation for the mass $M(r)$ as

$$\left\{ \frac{1}{4\pi r^2} \frac{dM(r)}{dr} \right\}^{n-2} \frac{d}{dr} \left\{ \frac{1}{4\pi r^2} \frac{dM(r)}{dr} \right\} = -\frac{G\rho_0^n}{\kappa_0} \cdot \frac{M(r)}{r^2}. \quad \dots\dots (10)$$

To solve this differential equation we take as a first approximation that $M(r) = 4\pi r^3 \bar{\rho}/3$ where $\bar{\rho}$ is the mean density of the whole sphere of radius R . Substituting this value of $M(r)$ in the second member and integrating we obtain as a second approximation that

$$M(r) = \frac{4}{3} \pi r^3 \rho_0 A^\alpha F\left\{-\alpha, \frac{3}{2}; \frac{5}{2}; \frac{Br^2}{A}\right\},$$

which gives equation (8) and no longer depends on the validity of (7).

Substituting back into the differential equation (10) we find that the left side is constant while the right side varies by a fraction $F\{-\alpha, 3/2; 5/2; BR^2/A\}$ of itself as r varies from 0 to R . This variation is seen to be equal to $\bar{\rho}/\rho_c$.

Now

$$1 \geq F\left\{-\alpha, \frac{3}{2}; \frac{5}{2}; \frac{BR^2}{A}\right\} \geq F\left\{-\alpha, \frac{3}{2}; \frac{5}{2}; 1\right\},$$

the lower bound being attained in the limit as $\bar{\rho} \rightarrow \infty$. Thus we see that, provided $F\{-\alpha, 3/2; 5/2; 1\}$ does not differ greatly from unity, the approximate solution (8) is valid for all values of $\bar{\rho}$. But

$$F\{-\alpha, 3/2; 5/2; 1\} = \frac{\Gamma(5/2)\Gamma(1+\alpha)}{\Gamma(5/2+\alpha)},$$

which is precisely unity when $\alpha=0$ and then decreases as α increases. Thus for α not too large the equation (8) is valid no matter how great is the variation of density throughout the sphere.

Combining (10) with the equation

$$\frac{3}{2r^2} \frac{dI(r)}{dr} = \frac{dM(r)}{dr},$$

where $I(r)$ is the moment of inertia of the sphere of radius r , we easily derive the integro-differential equation

$$\left\{ \frac{3}{8\pi r^4} \frac{dI(r)}{dr} \right\}^{n-2} \frac{d}{dr} \left\{ \frac{3}{8\pi r^4} \frac{dI(r)}{dr} \right\} = -\frac{G\rho_0^n}{\kappa_0} \frac{1}{r^2} \int_0^r \frac{3}{2r^2} \frac{dI(r)}{dr} dr,$$

and, on solving approximately by taking $I(r) = 8\pi r^5 \bar{\rho}/15$ in the second member, we find that

$$I(r) = \frac{8}{15} \pi \rho_0 r^5 A^\alpha F\left\{-\alpha, \frac{5}{2}; \frac{7}{2}; \frac{Br^2}{A}\right\},$$

which gives equation (9). On substituting back into the integro-differential equation we find that the left side is constant while the right side varies by the fraction $F\{-\alpha, 3/2; 5/2; BR^2/A\}$ of itself as r varies from 0 to R and so (9) is valid under the same conditions as (8).

V. AN ALTERNATE METHOD OF SOLUTION OF EQUATION (3)

As an alternative to the approximate method of Section IV we may take $\rho = H - Kr^2$ as a first approximation to the density in the second member of equation (3). On integrating and inserting the boundary conditions that ρ is finite at the origin, and $\rho = \rho_0$ when $r = R$ we have

$$\left(\frac{\rho}{\rho_0}\right)^{n-1} = 1 + \frac{4\pi G \rho_0(n-1)}{\kappa_0} \left\{ \frac{H}{6}(R^2 - r^2) - \frac{K}{20}(R^4 - r^4) \right\}. \quad \dots (11)$$

If we take

$$H - Kr^2 = \rho_c - \frac{\rho_c - \rho_0}{R^2} r^2, \quad \dots \dots \dots (12)$$

and substitute (11) back into the differential equation (3) we find that both sides of the equation vary from ρ_c to ρ_0 as r varies from 0 to R so that the solution (11) with the values of H and K given by (12) should be a valid approximation for compression of any magnitude.

Substituting the values of H and K given by equation (12) into (11) we see that, if ρ_c is large, then the density throughout the sphere is given asymptotically by

$$\frac{\rho}{\rho_0} \sim \left\{ \frac{4\pi G \rho_0}{\kappa_0 \alpha} \cdot \rho_c \left(\frac{R^2 - r^2}{6} - \frac{R^4 - r^4}{20R^2} \right) \right\}^\alpha \quad \dots\dots\dots (13)$$

VI. THE EXACT SOLUTION OF EQUATION (3) FOR $dx/dp=6/5$

When $n=6/5$ the equation (3) permits of an exact solution (Chandrasekhar 1939), namely,

$$\frac{4}{5}\pi G \frac{\rho_0^{6/5} \rho^{4/5}}{\kappa_0} = \frac{3C^2}{(1+C^2r^2)^2},$$

where C is an arbitrary constant. Inserting the boundary conditions $\rho=\rho_0$ when $r=R$ and $\rho=\rho_c$ when $r=0$ we obtain the relation between ρ_c and R

$$1 + \frac{4\pi G \rho_0^{6/5} \rho_c^{4/5}}{15\kappa_0} R^2 = \left(\frac{\rho_c}{\rho_0} \right)^{2/5} \quad \dots\dots\dots (14)$$

VII. DEDUCTIONS

(a) From the equation $A=1+BR^2$ combined with (8) we have

$$R^2 = \frac{A-1}{B} = \frac{1}{B} \left[\left(\frac{\bar{\rho}}{\rho_0 F\{-\alpha, 3/2; 5/2; BR^2/A\}} \right)^{1/\alpha} - 1 \right], \quad \dots\dots\dots (15)$$

so that in the limit as $\bar{\rho}$ increases indefinitely, we see that R^2 behaves like ρ^{n-2} . Hence in the limit as $\bar{\rho} \rightarrow \infty$ we have that for $n>2$, $R \rightarrow \infty$ and for $n<2$, $R \rightarrow 0$. Thus the case $n=2$ divides discontinuously the two types of behaviour.

When $n=2$, equation (15) reduces to

$$R^2 = \frac{1}{B} \left[\frac{\bar{\rho}}{\rho_0 F\{-1, 3/2; 5/2; BR^2/A\}} - 1 \right],$$

so that on taking the limit as $\bar{\rho} \rightarrow \infty$ we find that

$$R_{\max.} = \frac{1}{2\rho_0 \sqrt{\frac{15\kappa_0}{\pi G}}}.$$

On considering the exact solution (5) for $n=2$, we see that $\bar{\rho} \rightarrow \infty$ as $\beta R \rightarrow \pi$, so that the exact value for the upper bound of the radius is given by

$$R_{\max.} = \frac{\pi}{\beta} = \frac{1}{2\rho_0 \sqrt{\frac{\pi\kappa_0}{G}}}.$$

In Section IV the validity of the approximate equation (8) in the case of large $\bar{\rho}$ was shown to depend on α , so that the discrepancy between the exact

and approximate values of R_{\max} for $\alpha=1$, found in this section, gives an idea of the error involved.

As an alternative to the above we see from (13) that as $\rho_c \rightarrow \infty$, ρ_c behaves like $\rho_c^\alpha R^{2\alpha}$, or $R^{2\alpha}$ behaves like $\rho_c^{1-\alpha}$. Hence the results deduced from Section IV are supported by the solution of Section V.

(b) On combining equations (8) and (9) we obtain

$$a = \frac{2}{5} \frac{F\{-\alpha, 5/2; 7/2; BR^2/A\}}{F\{-\alpha, 3/2; 5/2; BR^2/A\}},$$

so that as $\bar{\rho} \rightarrow \infty$ we see that

$$a_{\min.} = \frac{2}{2\alpha+5}. \quad \dots\dots\dots (16)$$

From the exact solution (6) in the case $d\kappa/d\rho=2$ we see that, since $\beta R \rightarrow \pi$ as $\bar{\rho} \rightarrow \infty$,

$$a_{\min.} = \frac{2}{3} \left(\frac{\pi^2 - 6}{\pi^2} \right). \quad \dots\dots\dots (17)$$

It emerges therefore that the ratio I/MR^2 has a lower limit as the mean density of the sphere increases indefinitely and the approximate value $a_{\min.}=2/(2\alpha+5)$ is not far in error.

Using equation (13) and the integrals for mass and moment of inertia

$$a = \frac{I}{MR^2} = \frac{2}{3R^2} \frac{\int_0^R r^4 \rho(r) dr}{\int_0^R r^2 \rho(r) dr},$$

or

$$\begin{aligned} \lim_{\rho_c \rightarrow \infty} a &= \frac{2}{3R^2} \frac{\int_0^R r^4 \{(R^2 - r^2)/6 - (R^4 - r^4)/20R^2\}^\alpha dr}{\int_0^R r^2 \{(R^2 - r^2)/6 - (R^4 - r^4)/20R^2\}^\alpha dr} \\ &= \frac{2}{3} \frac{\int_0^1 u^4 \{(1 - u^2)(7 - 3u^2)\}^\alpha du}{\int_0^1 u^2 \{(1 - u^2)(7 - 3u^2)\}^\alpha du}, \end{aligned}$$

where we have made the substitution $u=r/R$.

Towards solving the integral

$$\int_0^1 u^4 (1 - u^2)^\alpha (7 - 3u^2)^\alpha du$$

we make the substitution $u^2=x$, so that it reduces to

$$\frac{1}{2} 7^\alpha \int_0^1 x^{3/2} (1-x)^\alpha \left(1 - \frac{3}{7}x\right)^\alpha dx,$$

which can be evaluated by expanding the last term in the integrand by the binomial theorem. On integrating term by term, the resulting series can be shown to be equal to

$$\frac{1}{2}7^\alpha \frac{\Gamma(\alpha+1)\Gamma(5/2)}{\Gamma(7/2+\alpha)} \cdot F\{-\alpha, 5/2; 7/2+\alpha; 3/7\}.$$

Similarly the integral

$$\int_0^1 u^2(1-u^2)^\alpha(7-3u^2)^\alpha du$$

can be shown to equal

$$\frac{1}{2}7^\alpha \frac{\Gamma(\alpha+1)\Gamma(3/2)}{\Gamma(5/2+\alpha)} \cdot F\{-\alpha, 3/2; 5/2+\alpha; 3/7\}.$$

Hence

$$\lim_{\rho_c \rightarrow \infty} a = \frac{2}{2\alpha+5} \frac{F\{-\alpha, 5/2; 7/2+\alpha; 3/7\}}{F\{-\alpha, 3/2; 5/2+\alpha; 3/7\}}, \quad \dots\dots (18)$$

which, when $\alpha=1$, is seen to be very close to the exact value of equation (17). It would appear from this agreement that the equation (18) gives a slightly better approximation to the value of $\lim_{\rho_c \rightarrow \infty} a$, for large α , than does equation (16).

In the case $\alpha=3/2$, corresponding to $dx/dp=5/3$, which is of importance, the equation (16) gives $\lim_{\rho_c \rightarrow \infty} a=0.25$, while the equation (18) gives $\lim_{\rho_c \rightarrow \infty} a=0.23$.

We thus take $a=0.23$ to be the lower bound of the ratio I/MR^2 for a homogeneous sphere when $dx/dp=5/3$, but note that the simple equation (16) gives a reasonably accurate approximation.

(c) Using equation (15) we can express the mass of the sphere in the form

$$M = \frac{4}{3}\pi \bar{\rho} \frac{1}{B^{3/2}} \left[\left(\frac{\bar{\rho}}{\rho_0 F\{-\alpha, 3/2; 5/2; BR^2/A\}} \right)^{1/\alpha} - 1 \right]^{3/2},$$

so that in the limit as $\bar{\rho}$ increases indefinitely we see that M behaves like $\rho^{3n/2-2}$, and thus M tends to a finite limit as $\bar{\rho} \rightarrow \infty$ if $n=4/3$.

This behaviour of M can also be deduced from equation (13) since for large values of the central density

$$\lim_{\rho_c \rightarrow \infty} M \rho_c^{-\alpha} = 4\pi \rho_0 \int_0^R r^2 \left[\frac{4\pi G \rho_0}{\kappa_0 \alpha} \left\{ \frac{R^2 - r^2}{6} - \frac{R^4 - r^4}{20R^2} \right\} \right]^\alpha dr,$$

or, on making the substitution $u=r/R$,

$$\begin{aligned} \lim_{\rho_c \rightarrow \infty} M \rho_c^{-\alpha} &= \frac{\pi \rho_0}{15} \left(\frac{4\pi G \rho_0}{\kappa_0 \alpha} \right)^\alpha R^{2\alpha+3} \int_0^1 u^2 \{ (1-u^2)(7-3u^2) \}^\alpha du \\ &= \frac{\pi \rho_0}{15} \left(\frac{4\pi G \rho_0}{\kappa_0 \alpha} \right)^\alpha R^{2\alpha+3} \frac{1}{2} 7^\alpha \frac{\Gamma(\alpha+1)\Gamma(3/2)}{\Gamma(5/2+\alpha)} F\{-\alpha, 3/2; 5/2+\alpha; 3/7\}. \end{aligned}$$

Hence as $\rho_c \rightarrow \infty$, M behaves like $\rho_c^\alpha R^{2\alpha+3}$.

On making use of the results of deduction (a) we see that for large central density M behaves like $\rho_c^{3n/2-2}$ and so M is finite if $n=4/3$.

This result agrees with that of Chandrasekhar (1939) who found an upper limit to the mass of a polytrope in the case where the pressure-density relation is of the form $p=A\rho^{4/3}$ corresponding to a completely degenerate electron gas when relativistic effects are taken into account.

It also follows from Chandrasekhar's treatment that values of $n<4/3$ are physically impossible; and it would appear that on the Thomas-Fermi theory no sphere could have an infinite mass since the ultimate pressure density relation for all substances is $p=A\rho^{4/3}$.

(d) The exact solution (14) for the case $d\kappa/dp=6/5$ bears out the previous result that $R\rightarrow 0$ as $\rho_c\rightarrow\infty$ if $d\kappa/dp<2$. In this case we can find the maximum radius which, since it corresponds to $dR/d\rho_c=0$, is

$$R_{\max.} = \sqrt{\left(\frac{15\kappa_0}{16\pi G\rho_0^2}\right)},$$

which occurs when $(\rho_c/\rho_0)^{2/5}=2$.

(e) Taking R as half its limiting value, i.e. $\beta R=\frac{1}{2}\pi$ in the exact solution for $d\kappa/dp=2$, we obtain from equations (5), (6), (4), and (1) that

$$\begin{aligned}\frac{\bar{\rho}}{\rho_0} &= \frac{12}{\pi^2}, \\ a &= 2 - \frac{16}{\pi^2}, \\ \frac{\rho_c}{\rho_0} &= \frac{\pi}{2}, \\ p_c &= \frac{\kappa_0}{2} \left\{ \frac{\pi^2}{4} - 1 \right\},\end{aligned}$$

where ρ_c and p_c are the central values of the density and pressure respectively. It is to be noted that a sphere of this radius shows very little compression.

(f) Birch (1952) and Keane (1954) have shown that, at least for the alkali metals, $d\kappa/dp$ is a decreasing function of pressure, so that, on assuming its universal applicability and in view of the quantum mechanical result for a completely degenerate electron gas (neglecting relativistic effects) that $d\kappa/dp\rightarrow 5/3$ as $p\rightarrow\infty$, it emerges from the foregoing results that no homogeneous elastic sphere can have the ratio I/MR^2 less than approximately 0.23, this value being approached as the mean density tends to infinity.

It is also evident from the previous results that, since $d\kappa/dp\rightarrow 5/3$ as $p\rightarrow\infty$, then there is an upper limit to the radius of the sphere of the order of $\sqrt{(\kappa_0/G\rho_0^2)}$. On taking $\kappa_0=10^{12}$ dyn/cm² and $\rho_0=1$ g/cm³ this upper limit is of the order of 10^4 km.

(g) Considering $\bar{\rho}\rightarrow\infty$ is a mathematical idealism and overlooks possible large-scale changes of interaction between the individual atoms. Clearly an infinite density is impossible and so some mechanism of destruction must be brought into operation before the limit is reached. It follows that the lower bound for the ratio I/MR^2 is unattainable.

(h) For the major planets the empirical values of I/MR^2 are all close to the minimum, 0.23, being, according to Ramsey (1951), for Jupiter 0.25, Saturn 0.22, Uranus 0.24, and Neptune 0.27. These low values suggest a certain amount of central concentration of heavy material, especially in Saturn, and also indicate that the radii of the major planets, all of the order of 10^4 km, are near the maximum for the particular materials of which they are composed.

(i) Ramsey (1951) has calculated the density distribution for Jupiter and Saturn on the assumption that they are composed of a homogeneous mixture of hydrogen and helium. From these calculations it was found that $a=0.28$ for Jupiter and has a slightly higher value for Saturn. These values are within the limits, for a homogeneous sphere, imposed by this paper but are so close to the lower limit as to suggest very forcibly that Jupiter is nearly the largest planet possible. In this connexion it should be borne in mind that, since we are considering $dx/dp \rightarrow 5/3$ as $p \rightarrow \infty$, the lower limit of a does not correspond to the maximum radius, but to the zero radius which is attained on shrinking of the sphere after the maximum radius has been reached. Hence the maximum radius corresponds to a value of a greater than 0.23.

It is difficult to ascertain ρ_0 and κ_0 for hydrogen since for a sphere of any size the hydrogen atoms are likely to form a metallic lattice. However, the value of 85,000 km shown by Ramsey (1950), using quantum mechanical calculations, to be the maximum radius of a hydrogen planet is not in conflict with the upper bound found in this paper.

VIII. REFERENCES

- BIRCH, F. (1939).—The variation of seismic velocities within a simplified earth model, in accordance with the theory of finite strain. *Bull. Seism. Soc. Amer.* **29**: 463-79.
- BIRCH, F. (1952).—Elasticity and constitution of the earth's interior. *J. Geophys. Res.* **57**: 227-86.
- CHANDRASEKHAR, S. (1939).—"An Introduction to the Study of Stellar Structure," (Univ. Chicago Press.)
- KEANE, A. (1954).—An investigation of finite strain in an isotropic material subjected to hydrostatic pressure and its seismological applications. *Aust. J. Phys.* **7**: 322-33.
- RAMSEY, W. H. (1950).—The planets and the white dwarfs. *Mon. Not. R. Astr. Soc.* **110**: 444-54.
- RAMSEY, W. H. (1951).—On the constitution of the major planets. *Mon. Not. R. Astr. Soc.* **111**: 427-47.

STRESS RELAXATION OF WOOL FIBRES IN WATER AT STRAINS OF 5-20 PER CENT.

By B. J. RIGBY*

[*Manuscript received July 5, 1954*]

Summary

A study has been made of the stress relaxation at fixed longitudinal strains of wool fibres immersed in water, in the post-yield region of their load-strain curves. The fibres were extended at controlled rates up to predetermined strains, particular attention being paid to the first few minutes of the subsequent stress relaxation. It was found that the relaxation curves so obtained could be represented by a sum of one, two, or three negative exponentials, which could, in general, be simplified into two main processes, (a) a relatively fast relaxation of stress occurring in about the first 60 sec and (b) a much slower relaxation. The activation energy of the fast process was found to be approximately 5 kcal/mol. Attempts to calculate activation energies for the slow process showed that the term corresponding to the energy was temperature dependent. Over the range 10-50 °C, the values of this term were found to increase from about 22 to 27 kcal/mol. The fast and slow processes are respectively ascribed to the breakdown of weaker and stronger types of cross-linking bonds in the keratin.

I. INTRODUCTION

When a keratin fibre is held at a fixed elongation under constant atmospheric conditions the stress decays, appearing to approach an asymptotic value.

According to the generally accepted picture of keratin, it is made up of polypeptide chains joined by a number of different cross-links. The cross-links in keratin include the strong cystine links, involving covalent disulphide bonds, and much weaker links involving hydrogen bonds, van der Waals forces, and electrostatic salt bonds. Speakman and Shah (1941) have attributed stress relaxation in keratin fibres to the breakdown of these cross-links under large strains.

Many conditions affect the rate and amount of stress relaxation, and a number of workers (Speakman 1928; Speakman and Shah 1941; Bull 1945; Katz and Tobolsky 1950) have examined the effect of some of these such as relative humidity, fibre extension, and temperature. Chemical pretreatment aimed at destroying disulphide bonds before relaxation has been carried out (Speakman 1928; Speakman and Shah 1941; Katz and Tobolsky 1950). Further, Katz and Tobolsky (1950) allowed the relaxation rate of an untreated fibre in water to become relatively slow before adding chemicals reagents to break these bonds.

The results of all these studies seem to indicate that the rate of relaxation is accelerated by any process which facilitates bond breakdown.

* Wool Textile Research Laboratories, C.S.I.R.O., Physics and Engineering Unit, Ryde, N.S.W.

No general equation has been found to fit the experimental curves, although a number of stress-time relationships have been reported which describe parts of them. Little work seems to have been done on the first few minutes of the stress decay, during which time there is a relatively rapid relaxation.

It appears in all of the published work that very rapid strain rates have been employed. The aim seems to have been to extend the fibres as rapidly as possible so as to minimize the relaxation taking place during the straining operation. No stress-strain curves at these high rates seem to have been recorded. Thus, strains, as reported in previous work, cannot be correlated as yet with the well-known regions of the stress-strain diagram, viz. "Hookean", "post-yield", etc. It was thought that useful information could be obtained by examining the phenomenon of stress relaxation in terms of these well-defined regions.

The strains in this work correspond to the post-yield region of the curve, with approximate limits of 5 and 20 per cent. in water. These strains were produced by known strain rates, which then became another parameter, not considered in previous work. Further, the effects of the magnitude of strain and temperature have been investigated with particular reference to the first minute of the relaxation process.

II. EXPERIMENTAL

In previous work various methods have been adopted to allow for the statistical distribution encountered in the measurement of fibre properties. Bull (1945), for example, used test specimens made up from about 20 fibres. Speakman and Shah (1941) and Katz and Tobolsky (1950) employed half-tension times and relative tensions in order to get a good degree of fibre to fibre reproducibility.

The procedure which has been used in this work is to obtain measurements on a single fibre subjected to various experimental conditions. This was done by immersing the fibre in water at 52 °C for 60 min, after which repetitive relaxation curves could be obtained. The experimental set-up consisted of a Cambridge extensometer as an agent for producing various strains and rates of strain and a Satham electrical resistance strain gauge, galvanometer, lamp, and scale for measuring load. The period of the galvanometer was approximately 2 sec; and it was found possible to read load correct to 0.04 g wt. in the early part of the relaxation, while for the slower portion the load readings were much more accurate. An examination shows that these limitations do not significantly affect the general results given in this paper. The strains were controlled by an automatic switch which could be preset at any desired value.

All the tests were carried out using wool from Corriedale sheep, hand fed for uniform growth. The coefficient of variation of the cross-sectional area along individual fibres was about 5 per cent., the average unstretched diameter, in water, being about 50 μ . A typical curve is illustrated in Figure 1 which shows the initial load-strain curve (obtained at 10 per cent./min) together with the subsequent relaxation of stress from a fixed strain of 20 per cent. The quantity, rate of straining, has been used in this work rather than "time of straining". The decrimped fibre length was standardized at 40 mm.

Most of the data are for a time range up to 200 sec (not including the time taken to extend the fibre), although some tests had time ranges up to 10 min. All the experiments were carried out with the fibre immersed in distilled water, the temperature of which was maintained constant to within $\pm 0.1^\circ\text{C}$.

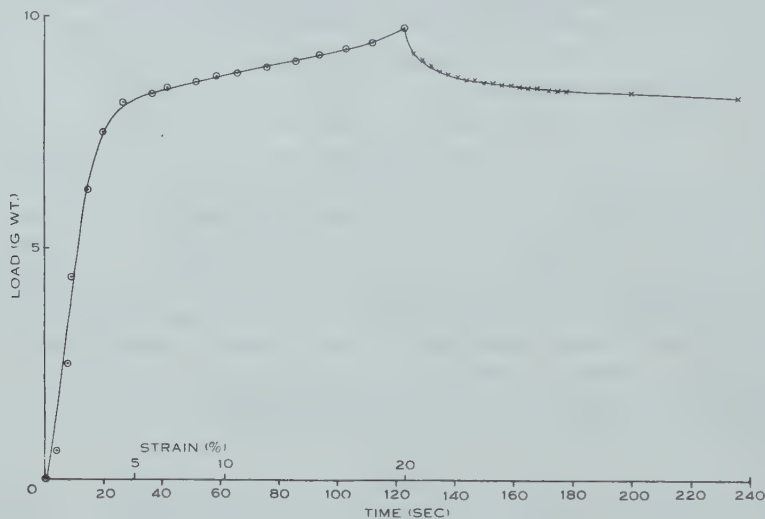


Fig. 1.—A typical curve, showing the initial load-strain curve (obtained at 10 per cent./min) together with the subsequent relaxation of stress from a fixed strain of 20 per cent.

III. RESULTS

Typical experimental results showing the effect of variation in each of three parameters are given in Figures 2, 3, and 4.

Each of these sets represents the effect upon the relaxation curve of the *variation* of only *one parameter* on a particular fibre. At least five different fibres were examined with each parameter, similar results being obtained for each. The range for each of the variables was as follows :

Strain, 5–20 per cent.

Strain rate, 0.75–250 per cent./min

Temperature, 5–50 $^\circ\text{C}$

Analysis of the relaxation curves obtained, over the time range 0–200 sec, showed that the curves could be represented by a sum of one, two, or three exponential terms, the number depending upon the particular experimental conditions, that is

$$S_t = \sum_i S_i \exp(-g_i t) + S_\infty, \quad \dots \dots \dots (1)$$

where S_t = stress at time t , S_∞ = asymptotic stress, S_i, g_i are functions of the parameters, $i=1, 2$, or 3.

It should be pointed out that it is not known, in analysing into exponentials, whether the expansion is unique, and also that over limited ranges one mathematical expression often closely approximates another. For example, in the

present instance each relaxation curve could be closely fitted by a rectangular hyperbolic law for about the first 60 sec. For these reasons it was not considered advisable to determine each S_i and g_i for a given curve. A significant feature

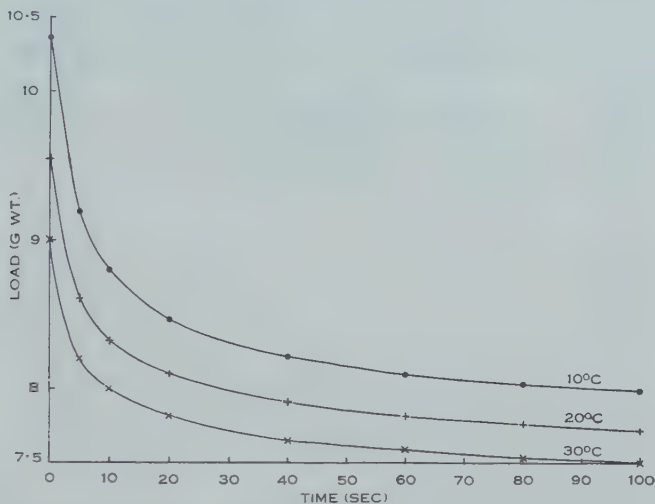


Fig. 2.—A typical set, showing the effect of the variation of temperature on the stress relaxation of a fibre which has been strained at a rate of 55 per cent./min to an extension of 20 per cent.

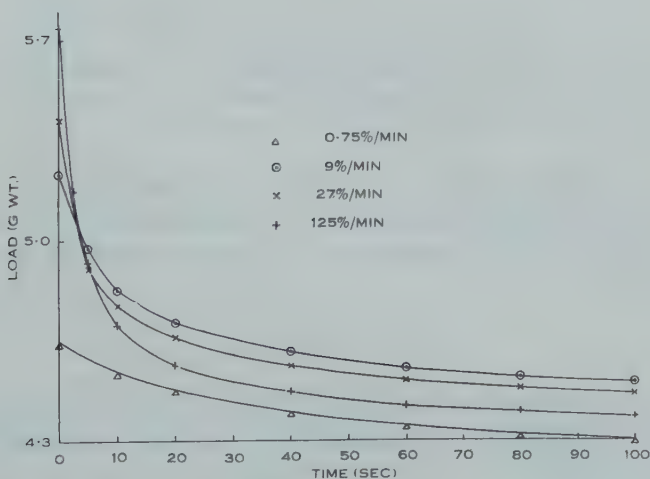


Fig. 3.—A typical set, showing the effect of the variation of strain rate upon stress relaxation. The temperature and total strain were maintained at 20 °C and 12.5 per cent. respectively.

of the present results, which will be discussed fully later, was that it was found possible to separate each curve into a *rapid* process of decay lasting about 40–60 sec followed by a considerably slower one (see Figs. 2, 3, and 4).

The effect of each parameter (the other two being kept constant) on the properties of the relaxation curve was as follows.

(a) *Relaxation as a Function of Initial Strain*

Examination of Figure 4, a typical set of curves, shows that :

(i) Increasing the initial strain brought about a proportional increase in the initial stress.

(ii) The relaxation curves were identical in shape, i.e. for any given period of time the stress decayed by the same amount.

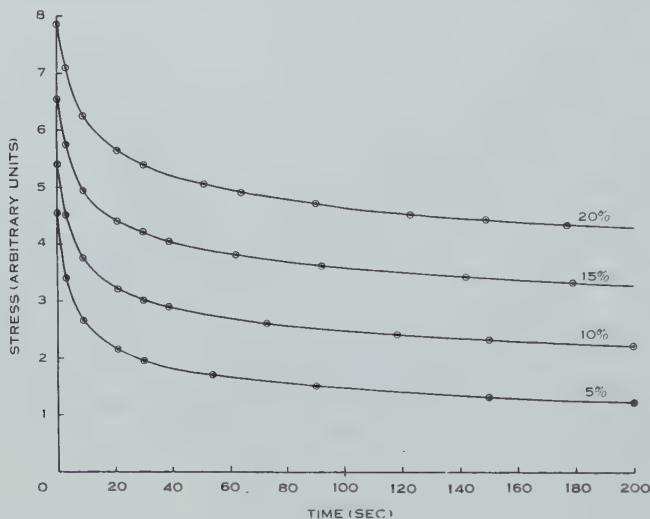


Fig. 4.—A set of curves showing the effect of the variation of total strain on the subsequent stress relaxation. The strain rate was 10 per cent./min and the temperature 19 °C.

(b) *Relaxation as a Function of Strain Rate*

(i) Referring to Figures 3 and 4, both the initial stress S_0 and $S_0 - S_t$ were found to vary linearly with $\ln \rho$ where ρ is the rate of straining up to the fixed strain (see Fig. 5).

$$S_0 = a \ln \rho + C_1, \quad \dots \dots \dots (2)$$

$$S_0 - S_t = a \ln \rho + C_2. \quad \dots \dots \dots (3)$$

(ii) The slope a of the lines (2) and (3) has a value in agreement with that predicted by a recent theory of viscoelasticity by Burte and Halsey (1947).

(iii) Analysis of one set of curves (Fig. 3) into exponentials, showed that the curve produced by the slowest rate of straining (0.75 per cent./min) reduced to a single exponential term. The curves due to the higher rates appeared to have this basic component plus others.

(c) *Relaxation as a Function of Temperature (see Fig. 2)*

(i) The initial stress S_0 was found to decrease almost uniformly with an increase in temperature.

(ii) As mentioned earlier, the relaxation curves can be separated into an initial rapid process, followed by a much slower one. Activation energies were calculated for each of these processes in the following manner (Glasstone, Laidler, and Eyring 1941).

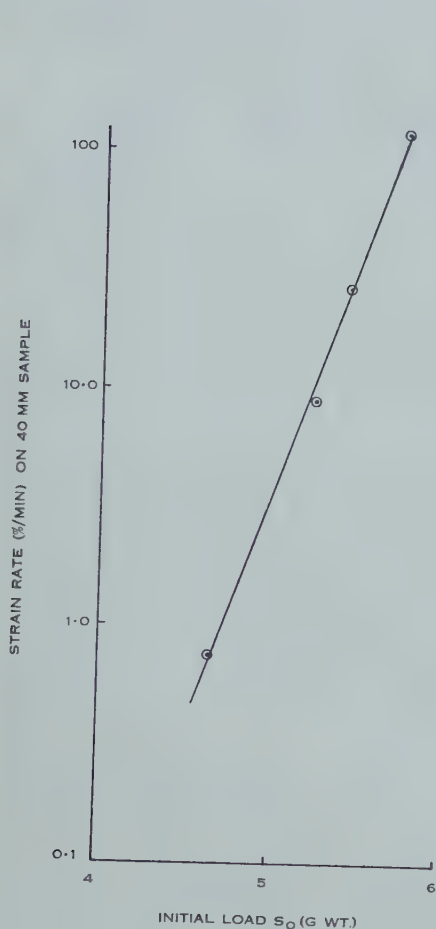


Fig. 5.

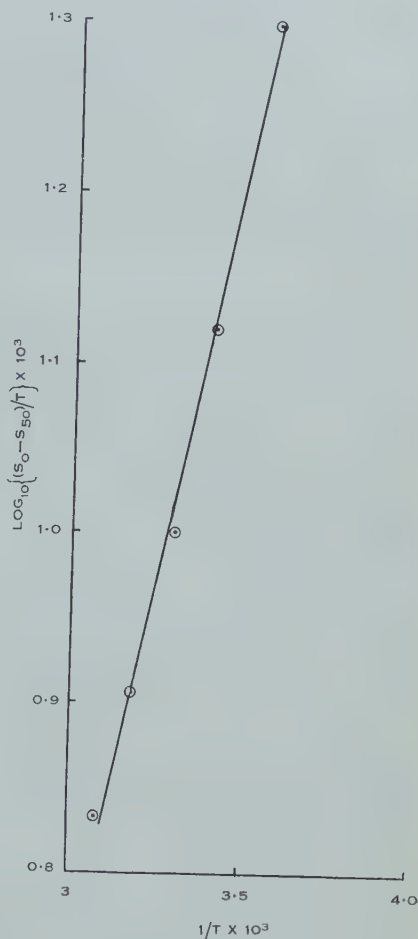


Fig. 6.

Fig. 5.—A plot of the initial stress S_0 , against the logarithm of the strain rate, with the temperature and total strain held at 20 °C and 12.5 per cent. respectively.

Fig. 6.—A plot of $\log_{10}\{(S_0 - S_{50})/T\}$ against $1/T$. The slope of this line is a measure of the activation energy of the fast relaxation process (see text).

The fast process.—It was assumed that the amount of stress decay in a given time was proportional to the specific rate constant (k') for the process ;

$$k' = \frac{kT}{h} \exp(-E_1/RT),$$

that is,

$$S_0 - S_t \propto T \exp(-E_1/RT),$$

where S_0 and S_t are defined above, T is the absolute temperature, k is Boltzmann's constant, h is Planck's constant, E_1 is the activation energy for the process, R is the universal gas constant.

Thus a plot of $\ln \{(S_0 - S_t)/T\}$ against $1/T$ should give a straight line, with a slope equal to E_1/R . Five fibres were studied in the temperature range 5–50 °C and straight lines were in fact obtained. A typical plot is shown in Figure 6. These straight lines were plotted for values of time up to 50 sec since the fast process is practically completed by this time. The value of E_1 calculated from these lines varied from 4 to 7 kcal/mol, the average value being 5 kcal/mol.

The slow process.—The stress decay over a long period of time is predominantly due to the slower process, and in this case the time t' taken for the stress to fall to a given value should be inversely proportional to the rate constant (k').

Thus $t'T \propto \exp(-E_2/RT)$ where E_2 is the activation energy for the process.

A plot of $\ln(t'T)$ against $1/T$ should give a straight line with a slope equal to E_2/R . Seven fibres were studied over the range 10–50 °C, but the value of E_2 was found to be temperature dependent, varying from 22 to 27 kcal/mol.

IV. DISCUSSION OF RESULTS AND CONCLUSIONS

Most of the results can be interpreted using a simple extension of a suggestion made by Speakman and Shah (1941). They suggested that the breaking of cross-linking bonds in keratin followed by molecular rearrangement is responsible for stress relaxation. An equation of the form (1) could, for instance, be reasoned in the following manner. Suppose that after a longitudinal strain has been produced on a fibre there will be a certain number of cross-links in a strained state. Some may have broken, and re-formed during the straining period—this number will depend upon the rate of straining, but whatever has happened during the extension it should only affect the amounts of stress decay and the initial stress in the stress relaxation test. This very simple picture explains the form of the relaxation curves by only requiring that at the completion of any fibre extension there are a number of strained cross-links. It seems reasonable to assume that the total stress energy which can relax will be proportional to the number of these strained bonds. These bonds will immediately begin to revert to their stable energy state and this energy decrease will be manifested by a stress relaxation. Now, if these bonds revert with a rate which is proportional to the total number n strained at any time, thus

$$\frac{dn}{dt} = -gn,$$

then, in view of the statement that the number n is proportional to the stress which can relax, we get upon integration,

$$S_t = \text{const.} \times \exp(-gt).$$

If there are a number of distinct cross-links, each having its own value of g , then

$$S_t = \sum_i S_i \exp(-g_i t).$$

In the equation given earlier, viz.

$$S_t = \sum_i S_i \exp(-g_i t) + S_\infty,$$

S_∞ would represent the stress energy contained in the main chains.

This exponential equation explains why the relaxation curves with strain as the only parameter are identical in shape but displaced along the stress axis. Bull (1945) shows this effect quite clearly graphically.

When the rate or time of straining is varied this simple theory would require that the amount of relaxable stress should increase with the strain rate ρ . This is verified by the data summarized in equations (2) and (3), but it would be difficult to derive equation (2) from the theory discussed above, because the stress S_0 is determined by other factors as well as the relaxation which occurs during straining.

Figure 3 is interesting for in this particular fibre it appears as though the slow process has been isolated. The bottom curve in this set can be represented by a single exponential term and, since it was obtained after a very slow rate of straining, it seems probable that the weak bond breakdown (which would give other exponential terms) has been in equilibrium with the molecular adjustments due to straining at all times. Further it will be noticed that the curve representing this slow process is parallel to the other three curves, after about 25 sec; indeed, analysis shows it to be a component of each of the four curves.

Because of the dependence upon temperature of E_2 there is strictly no justification in calling it an activation energy. However, the large difference between E_1 and E_2 supports the contention that the fast and slow processes are due to weaker and stronger bonds respectively. It may be noted that the value of 5 kcal/mol for E_1 could be ascribed to hydrogen bonds.

V. ACKNOWLEDGMENT

The author wishes to acknowledge the assistance of Mr. T. W. Mitchell in these experiments.

VI. REFERENCES

- BULL, H. B. (1945).—*J. Amer. Chem. Soc.* **67**: 533.
 BURTE, H., and HALSEY, G. (1947).—*Text. Res. J.* **17**: 465.
 GLASSTONE, S., LAIDLER, K. G., and EYRING, H. (1941).—"Theory of Rate Processes," 1st Ed. (McGraw-Hill: New York.)
 KATZ, J. M., and TOBOLSKY, A. V. (1950).—*Text. Res. J.* **20**: 87.
 SPEAKMAN, J. B. (1928).—*Proc. Roy. Soc. B* **103**: 377.
 SPEAKMAN, J. B., and SHAH, S. Y. (1941).—*J. Soc. Dy. Col.* **57**: 108.

SHORT COMMUNICATIONS

A PROPOSED MODIFICATION TO THE C.S.I.R.O. MARK I COMPUTER*

By B. E. SWIRE†

The order code of the C.S.I.R.O. Mark I computer, as described by Pearcey and Hill in their recent papers (1953*a*, 1953*b*, 1954), is already more extensive and flexible than that of any other machine of which a description has been published. It still offers some scope for expansion, however, and it is the purpose of this note to put forward one particular suggestion for the addition of an entirely new type of order. "General purpose" computers, so-called, are at present by no means so general in their application as users would desire, and broadly speaking there are two types of attack on this problem. On the one hand one can incorporate in the order code more and more special functions so as to increase the number of types of work with which the computer can directly deal. On the other hand one can seek new types of order which tend to increase the flexibility of a machine without special reference to any particular application. The suggestion put forward here is in line with this latter approach, which is believed to be the more valuable of the two.

The importance of flexibility has been recognized in the logical design of C.S.I.R.O. Mark I computer, and noteworthy success has been achieved by the use of an arithmetical unit containing a multiplicity of registers having addition, subtraction, and discrimination facilities (Beard and Pearcey 1952), and by the adoption of an order code in which each order is expressed as a transfer from a "source" to a "destination" each of which is separately and independently specified by a five-digit code. Programmes expressed in this code are very considerably shorter than programmes expressed in a more conventional one-address code such as that of the EDSAC. Nevertheless it has been found necessary to use interpretive programming techniques when it is desired to work to multiple-word accuracy, in floating-point arithmetic, or with complex numbers.

The interpretive programmes, or "hyper-programmes", as described by Pearcey and Hill, consist of a mixture of two types of command, namely, commands coded as ordinary machine commands and commands requiring interpretation; but for each command, whether it is one requiring interpretation or not, the interpretation routine of some 20 or 30 commands must be traversed. It is, in fact, the interpretation routine which distinguishes between a command

* Manuscript received October 19, 1954.

† Aeronautical Research Laboratories, Department of Supply, Melbourne; present address: Adolph Basser Computing Laboratory, University of Sydney.

coded as a machine command, and a command labelled by the use of the null destination Z to indicate that it requires interpretation.

With this interpretive system, as with similar systems used on other machines, a gain in flexibility and convenience is achieved at a considerable cost in speed. But suppose now that a facility be built into the machine for distinguishing between machine commands and interpretive or "hyper" commands, in such a way that the former are performed directly, while, when one of the latter is encountered, control is directed to the head of the interpretation routine with simultaneous storage of the link datum to enable return of control to the next programme command after performance of the hyper-function. If such a facility were available the interpretation routine would be considerably shortened and furthermore would be used only when strictly necessary. A considerable gain in speed over the interpretive system described by Pearcey and Hill would be achieved. In fact, this facility, if available, would become the normal method of calling in sub-routines, and the distinction between direct and interpretive programming would vanish.

Not enough is known of the detailed design of existing overseas machines to say whether the proposed facility could be readily added to any of them. The design of the C.S.I.R.O. Mark I machine is, however, such as to allow its ready inclusion. As mentioned above, hyper-commands as at present used on the C.S.I.R.O. machine are labelled by the use of the Z destination code. If this Z code appears in a normal machine command, a null command results since no gates are opened. Suppose now that a small addition were made to the machine so that this Z destination did control certain gates so as to perform the following functions :

(a) read out the content of the sequence register in digit positions 11–20, leaving the sequence register clear ;

(b) read out the content of the lower half of the interpreter register in digit positions 1–10, leaving the address digits in the upper half of the interpreter undisturbed (as is the case with the $K+$ destination) ;

(c) open the read-in gate to register B so that the full word assembled in (a) and (b) is placed in register B .

Since the sequence register has been cleared, the next command executed would be that in location zero, which would be the head location of the interpretation routine. The first command in this routine would be $0 \rightarrow H_u$, whereupon the address digits previously left in the interpreter would be placed in the H register. All the data required are then available : the function code in digit positions 6–10 of B ; the address in H ; and the link datum in digit positions 11–20 of B . It would be a simple matter to design an interpretation routine to direct control via a directory to the appropriate sub-routine. The interpretation routine should preferably not use register B ; then each sub-routine could commence with the command $(B) \rightarrow x$, where x is a store location allocated within each sub-routine to hold the link datum. Each sub-routine would then conclude with the command $(x) \rightarrow S$ to return control to the next programme command (the single command $(B) \rightarrow S$ at the end would suffice for linking if

the sub-routine did not require to make use of register *B*). The present practice of using some of the *D* registers for holding link data would then be avoided, and these registers freed for other purposes.

The facility just proposed is not available on any existing computer, nor, so far as the author is aware, has it been previously suggested. There have been previous suggestions for calling in sub-routines by means of a single order. The Circle computer (Greig 1953), for example, includes what is known as the "Function Table" order; this does enable a sub-routine to be called in by a single command, but this command does not, and cannot, include an address or other parameter for use by the sub-routine called in. The function table order therefore does not have the character of a true hyper-command such as is possessed by the new type of order that has been here described.

The author acknowledges indebtedness to the Aeronautical Research Laboratories, Department of Supply, and the Radiophysics Laboratory, C.S.I.R.O., for jointly making it possible for him to become familiar with the C.S.I.R.O. Mark I computer and to carry out certain computing work with it. Particular appreciation is expressed of the unstinted help given by Mr. Pearcey and Mr. Beard and the members of their team at the Radiophysics Laboratory.

References

- BEARD, M., and PEARCEY, T. (1952).—*J. Sci. Instrum.* **29**: 305.
GREIG, J. (1953).—*Math. Tab., Wash.* **7**: 249.
PEARCEY, T., and HILL, G. W. (1953a).—*Aust. J. Phys.* **6**: 316.
PEARCEY, T., and HILL, G. W. (1953b).—*Aust. J. Phys.* **6**: 335.
PEARCEY, T., and HILL, G. W. (1954).—*Aust. J. Phys.* **7**: 485.

MULTIPLE IMAGES IN REFLECTING MICROSCOPES*

By W. H. STEEL†

Introduction

The normal image of a reflecting microscope is formed by a single reflection at each mirror surface. However, other images are possible; higher order images formed by two or more reflections at each surface, and a zero order "image" which is the object itself, if there is an unobstructed light path through the system. In a mirror system possessing central obscuration of the pupil, this unwanted light usually does not reach the axial image point, but it may give rise to a spread of light across the outer parts of the field, so reducing the contrast. The effects of these extra images are most noticeable when the objective is used in reverse as a condenser.

* Manuscript received July 1, 1954.

† Division of Physics, C.S.I.R.O., University Grounds, Sydney.

These false images have been indicated by Norris, Seeds, and Wilkins (1951) and they have shown stops placed to eliminate them. In the present paper, methods of choosing the position of these stops are discussed in more detail, reference being made to the consequent vignetting of the primary image; this follows a demonstration that it is impossible to secure a corrected secondary image having any advantages over the primary image.

Correction of the Secondary Image

The secondary image, formed by two reflections at each mirror, is studied by the approximate method used by the author (Steel 1953) for investigating the primary image of a monocentric mirror pair. The objective is considered in reverse so that its magnification $-g$, where $u_1 = -gu'_4$, is the reciprocal of the magnification when used as a microscope objective.

The third order spherical aberration ΣW_p is corrected when

$$\sigma = \frac{11 - 10g + 11g^2 \pm (1+g)\sqrt{(21 - 22g + 21g^2)}}{2(5 - 6g + 5g^2)}, \quad \dots \quad (1)$$

where σ is the ratio of the radii of the two mirrors,

$$\sigma = r_1/r_2.$$

The two "geometrical" obscurations, as defined by the above article, are

$$q_1 = (\sigma + 5g\sigma - 4g)/(2 - \sigma - 2g + 3g\sigma), \quad \dots \quad (2)$$

$$q'_2 = (5 - 4\sigma + g)/(3 - 2\sigma - g + 2g\sigma). \quad \dots \quad (3)$$

Object at Infinity.—The case of a microscope objective of high magnification can be studied approximately by considering $1/g$ to be infinite. Equation (1) then gives $\sigma = (11 \pm \sqrt{21})/10$.

As in the case of the primary image, the larger root, corresponding to a Cassegrain type objective, is of little practical interest as the image point is virtual. The smaller root, $\sigma = 0.6417$, corresponds to a Schwarzschild type objective for which $q_1 = 0.284$, $q'_2 = 1.417$. The physical obscuration p is given by $1/q'_2$ and is 0.706; for the primary image it is 0.447.

The fifth order spherical aberration is also found to be less favourable being 1.27 times that for the primary image at the same numerical aperture.

Finite Object.—The closest that the object can conveniently be placed to the objective is at the pole of the concave mirror. Then $q_1 = 0$ and it is found that $1/g = 3$ and $p = 0.6$. For the object at the same position, the primary image gives $1/g = 5$, $p = 0.333$.

Hence, for a given magnification, the secondary image has a higher obscuration than the primary image and its quality is therefore inferior.

Elimination of the Secondary Image

We now consider the elimination of the secondary image in a monocentric mirror pair, corrected for an object at infinity. The methods used are applicable to other types of reflecting systems.

Axial Rays.—The paths of the axial rays from infinity that form each image are shown in Figure 1, this figure corresponding to that given by Norris, Seeds, and Wilkins (1951). The rays forming the primary image are shown above the optical axis and those forming the secondary image below. When the system is used in the reverse sense as an objective, no axial rays from the object can form a secondary image.

In a monocentric objective, the obscuration for the secondary ray is 0.644 and it contains 0.146 of the light in the primary image. The secondary focal length is one-half the primary focal length and the secondary rays cross the primary field at angles above $0.288u$, where u is the numerical aperture of the objective. A field stop should therefore be used to limit the object to half this size.

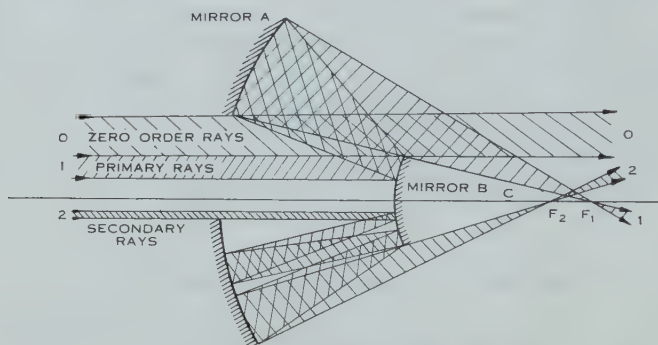


Fig. 1.—Axial rays forming primary and secondary images.

If no field stop can be used, or if a field greater than $0.144u$ is required, the secondary rays should be blocked completely. At their first reflection at both mirrors and also in the final image space, the axial rays forming the secondary and higher order images all lie inside the primary rays and so a central stop placed at any of these positions will eliminate them. Oblique rays, however, require further consideration.

Oblique Rays.—Zero order image. This effect is caused by direct light from the object reaching the image plane. It can occur in both directions, either when the system is used as an objective or as a condenser. However, in both cases, this light can be kept outside the primary image provided a field stop is used at the object or the image, in combination with a suitable choice of the size of the hole in the concave mirror.

For a monocentric system, it is found that the angular semi-field should be limited by a stop to a value less than $0.224u$. For non-concentric systems with obscuration p , the field would be $up/2$.

Secondary image. The effect on oblique rays of stops at the various positions suggested for axial rays can be seen in Figure 2. This shows the various apertures and obstructions presented by the system to an object at infinity, the scale being for a numerical aperture of 0.5 . The stops for the primary rays are shown above the optical axis and those for the secondary rays below. The lines at A and B represent the positions of the concave and convex mirrors

respectively, at A' and B' their first images and at A'' and B'' their second images in the whole system.

The full line at A and at its images represents the reflecting part of the concave mirror and the broken line the minimum size of the hole, both calculated for axial rays. To prevent vignetting for a field of semi-angle e , the radii of the hole and of the outer edge of the mirror should both be increased by $3.24ef'$. The radius of the inside limit of the reflecting surface should be decreased by the same amount except when this edge is used as a stop to cut out secondary rays. An upper limit can be placed on the size of the hole if it is to block the zero order image, so in general the hole should be smaller than the inside limit of the reflecting surface, the intervening portion being blackened.

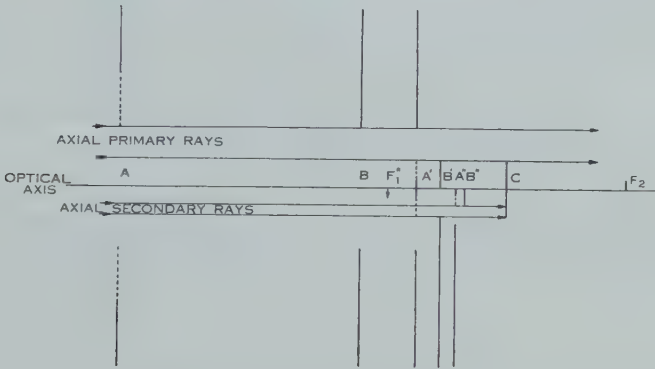


Fig. 2.—Obstructions and apertures for object at infinity.

The convex mirror cannot be increased in diameter beyond the value chosen for axial rays without increasing the obscuration. It acts as an aperture for primary rays at B and as an obscuration at B' . For secondary rays, B' acts as an aperture and rays that pass through it continue on to form the secondary image.

The primary image plane as seen by rays forming the secondary image is at F_1'' , the size shown being the maximum as indicated above. If a larger field is desired, no rays should be allowed to pass through the aperture at B' and this may be done by a central stop either at A' or B and another at C , the centre of curvature.

A single stop at the focus of the secondary rays is also effective. Its image for primary rays is at F_2' and the stop should be small enough not to vignette these rays.

Vignetting of the primary image. Provided that the hole and the outside edge of the concave mirror are made larger than the size demanded for axial rays by the amount indicated above, there is no vignetting at the edge of the pupil. But, if a central stop is used at the centre of curvature in combination with one at either of the mirrors to eliminate secondary rays, there are obstructions at C and B' and either at A' or B and some vignetting at the edge of the obstruction; it is smaller if the second stop is at the concave mirror.

If a stop at the secondary focus is used, there is no vignetting nor any light due to a secondary image for angular semi-fields less than $0.175u$. A stop in this position, however, would reduce the working distance of the microscope, and the gain in field size would be small.

Conclusions

The detailed discussion above has been based on paraxial optics and an object at infinity, but the author has obtained similar results for particular cases treated by trigonometrical ray traces. Similar considerations would also apply to non-concentric objectives and to objectives made of more than two mirrors, although in this latter case the treatment becomes more complicated, for every pair of mirrors can give rise to multiple reflections.

Thus it is most important to use a field stop at the object or image of all mirror systems. Auxiliary central stops are also advantageous in eliminating the light due to false images.

The author wishes to thank Mr. W. R. Blevin, of the Division of Physics, C.S.I.R.O., for drawing his attention to this problem.

References

- NORRIS, K. P., SEEDS, W. E., and WILKINS, M. H. F. (1951).—*J. Opt. Soc. Amer.* **41** : 111.
 STEEL, W. H. (1953).—*Rev. Opt. (Théor. Instrum.)* **32** : 269 (Thèse, Paris (1952)).

COSMIC RAYS AND AIR MASS EFFECTS AT MACQUARIE ISLAND*

By R. M. JACKLYN†

In a recent paper (Jacklyn 1954) it was shown that the barometer coefficient for cosmic rays at Macquarie Island varied consistently according to the type of air mass present over the station. This was found to be due to the different average values of $\delta H/\delta B$ (the rate of change of the height of the assumed production level for mesons with surface pressure) in the two air mass types considered. For temperate air $\delta H/\delta B$ was practically zero ($+0.0002$ per cent./mb)

TABLE 1
REGRESSION OF HEIGHT ON SURFACE PRESSURE

Pressure level (mb)	800	600	400	200	100
Polar air $\delta H/\delta B$ (km/mb $\times 10^2$)	+0.85	+0.975	+1.128	+1.189	+1.341
<i>r</i>	0.992	0.963	0.917	0.904	0.833
Temperate air $\delta H/\delta B$ (km/mb $\times 10^2$)	+0.762	+0.701	+0.640	+0.366	0
<i>r</i>	0.958	0.875	0.686	0.330	0.106

* Manuscript received August 30, 1954.

† Physics Department, University of Tasmania.

and for polar air had the value $+0.1028$ per cent./mb. Since then, an analysis of these air mass types has shown why one should expect the ratio to vary in the way it does.

For each type of air mass the total regression of height H on surface pressure B was calculated for the following pressure levels: 100, 200, 400, 600, and 800 mb, using data obtained from the daily radiosonde flights at Macquarie Island. The regression and correlation coefficients are set out in Table 1.

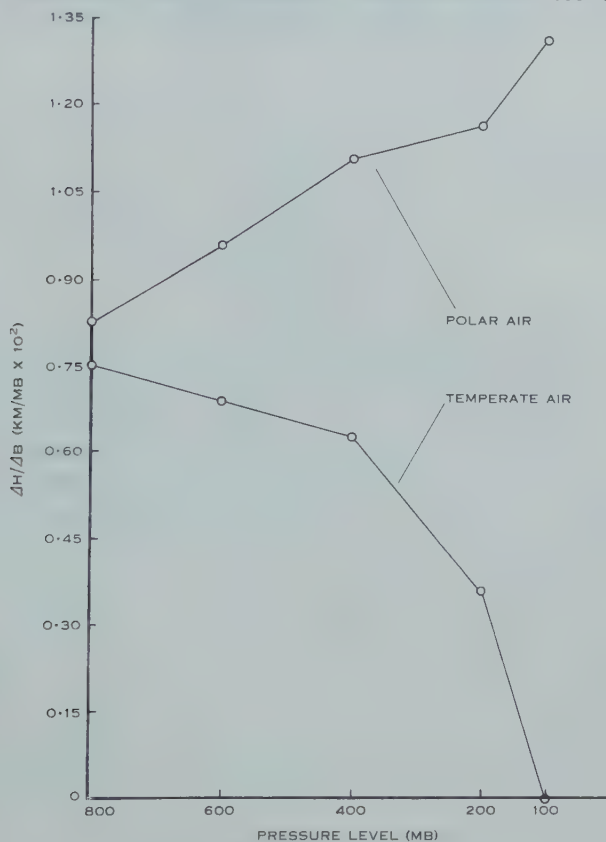


Fig. 1.—Variation of height H of pressure level per millibar increase of surface pressure B , for the series of pressure levels 800, 600, 400, 200, 100 mb, under polar and temperate air mass conditions.

From these, and from Figure 1, it can be seen that $\delta H/\delta B$ progressively increases as one proceeds upwards in polar air, and in temperate air it progressively decreases. In other words, there is a tendency for polar air to expand at all levels, and for temperate air to contract, as surface pressure increases.

If then, the height of production, H , of mesons, is considered as a function of the mean atmospheric temperature θ and the surface pressure B , and we have

$$\frac{dH}{dB} = \left(\frac{\partial H}{\partial \theta} \right)_B \frac{d\theta}{dB} + \left(\frac{\partial H}{\partial B} \right)_\theta,$$

it seems that $d\theta/dB$, the rate of expansion with surface pressure, would be positive in polar air and negative in temperate air. As a measure of the mean atmospheric temperature θ , the thickness of the atmosphere between 100 and 1000 mb was taken, and the regression coefficient $d\theta_P/dB$ and $d\theta_T/dB$ were found, the suffixes P and T referring to polar and temperate air respectively. The values were

$$\begin{aligned}\frac{d\theta_P}{dB} &= +0.521 \times 10^{-2} \text{ km/mB}, \\ \frac{d\theta_T}{dB} &= -0.808 \times 10^{-2} \text{ km/mB}.\end{aligned}$$

A check on the consistency of the figures can be made if it is assumed that $(\partial H/\partial\theta)_B$ and $(\partial H/\partial B)_\theta$ are constants; then

$$\frac{\delta H_P}{\delta B} - \frac{\delta H_T}{\delta B} = \left(\frac{\partial H}{\partial\theta}\right)_B \left(\frac{d\theta_P}{dB} - \frac{d\theta_T}{dB}\right).$$

Knowing the values of $\delta H_P/\delta B$ (approximately $+1.34 \times 10^{-2}$ km/mB) and $\delta H_T/\delta B$ (approximately 0 km/mB) arrived at from the radiosonde data independently of the temperature coefficients, $(\partial H/\partial\theta)_B$ is found to be 1.01, close to the expected value of unity (i.e. the method of measurement of θ implies that δH and $\delta\theta$ may each be considered as a measure of the variation of the depth of the atmosphere between 100 and 1000 mb at constant surface pressure).

It seems that these effects are characteristic of the pattern of weather associated with the passage of cold fronts at Macquarie Island during winter; that is, general atmospheric warming of temperate air precedes the arrival of the low, and is followed again by atmospheric warming of polar air accompanied by increasing surface pressures, after the front has passed. Presumably, this is true of other places where marked increases of cosmic ray intensities, "corrected" for pressure, have followed the passage of cold fronts. Since the Macquarie Island measurements agree substantially with those found by Loughridge and Gast (1940), using a shielded ionization chamber at sea off the west coast of North America, a similar maritime weather pattern probably held there. But Nishina *et al.* (1940), using an ionization chamber at Tokyo, found no significant effect for cold fronts, and this may have been due to the modifying effect of the neighbouring land on the atmospheric temperature characteristics of the air masses.

The author would like to thank Professor A. L. McAulay, Dr. A. G. Fenton, and Mr. N. R. Parsons for several helpful discussions, and Mrs. P. James for her assistance with the computations. The data used were obtained from the records of the Australian National Antarctic Research Expedition.

References

- JACKLYN, R. M. (1954).—*Aust. J. Phys.* **7**: 315–21.
 LOUGHRIDGE, D. H., and GAST, P. F. (1940).—*Phys. Rev.* **58**: 583–5.
 NISHINA, Y., SEKIDO, Y., SIMAMURA, H., and ARAKAWA, H. (1940).—*Phys. Rev.* **57**: 1050.

THE SAMPLING DISTRIBUTIONS OF STATISTICS DERIVED FROM THE MULTIVARIATE t -DISTRIBUTION

By E. A. CORNISH*

[*Manuscript received November 30, 1954*]

Summary

The sampling distributions of the more important statistical derivatives from the multivariate t -distribution are established.

I. INTRODUCTION

A multivariate generalization of Student's t -distribution has been considered by Dunnett and Sobel (1954) in connexion with certain multiple decision problems concerned with the ranking, according to their mean values, of normal populations having a common unknown variance. These authors studied the probability integral of the bivariate population in detail, giving exact and asymptotic expressions, and tables for certain special cases. The particular applications to which they refer have been discussed by Bechhofer, Dunnett, and Sobel (1954). The same general distribution was derived, and its principal properties established independently by the author (Cornish 1954), when considering the pretreatment to be given to certain types of replicated experimental observations, before applying tests of normality. This distribution possesses properties which make it suitable as a basis for exact tests of significance in various problems, and Dunnett and Sobel have taken the first step towards its use in practice by providing tables of the probability integral. In this paper, we shall be concerned with sampling distributions of statistics derived from the multivariate t -distribution. The general sampling distribution of the means and sums of squares and products of the variates is first established, and from it, the sampling distributions of the more important statistical derivatives are obtained.

II. GENERAL DISTRIBUTION OF MEANS AND SUMS OF SQUARES AND PRODUCTS

Suppose x_1, x_2, \dots, x_p are distributed in a non-singular multivariate normal distribution having a null vector of means and variance-covariance matrix $\sigma^2 \mathbf{R}$. In the class of cases to be considered, the symmetric correlation matrix $\mathbf{R} = [\rho_{ij}]$ is known, but the variance σ^2 is unknown. If s^2 is an estimate of σ^2 , based on ν degrees of freedom and distributed independently of x_1, x_2, \dots, x_p , then, as indicated above, it has been shown that the variates

$$t_i = x_i/s, \quad i=1, 2, \dots, p, \quad \dots\dots\dots (1)$$

* Division of Mathematical Statistics, C.S.I.R.O., Adelaide.

have the distribution

$$\frac{|\mathbf{R}|^{-\frac{1}{2}} \Gamma_{\frac{1}{2}}(\nu+p)}{(\pi\nu)^{\frac{1}{2}p} \Gamma_{\frac{1}{2}}^{\frac{1}{2}}\nu} (1 + \mathbf{t}' \mathbf{R}^{-1} \mathbf{t} / \nu)^{-\frac{1}{2}(\nu+p)} d\mathbf{t}, \quad \dots\dots\dots (2)$$

where \mathbf{t}' is the row vector $[t_1 \ t_2 \ \dots \ t_p]$.

If n samples are taken from the multivariate normal distribution yielding the variate values

$$x_{ik}, \quad \begin{matrix} i=1, 2, \dots, p, \\ k=1, 2, \dots, n, \end{matrix}$$

the probability of obtaining the corresponding sets of t -variates, defined by

$$t_{ik} = x_{ik}/s, \quad \dots\dots\dots (3)$$

is

$$\left\{ \frac{|\mathbf{R}|^{-\frac{1}{2}} \Gamma_{\frac{1}{2}}(\nu+p)}{(\pi\nu)^{\frac{1}{2}p} \Gamma_{\frac{1}{2}}^{\frac{1}{2}}\nu} \right\}^n \prod_{k=1}^n (1 + \mathbf{t}_k' \mathbf{R}^{-1} \mathbf{t}_k / \nu)^{-\frac{1}{2}(\nu+p)} d\mathbf{t}_k. \quad \dots\dots (4)$$

The means and sums of squares and products of deviations of these t -variates are defined respectively by the relations

$$\left. \begin{aligned} \bar{t}_i &= \frac{1}{n} \sum_{k=1}^n t_{ik}, & i &= 1, 2, \dots, p, \\ T_{ij} &= \sum_{k=1}^n (t_{ik} - \bar{t}_i)(t_{jk} - \bar{t}_j), & i, j &= 1, 2, \dots, p. \end{aligned} \right\} \quad \dots\dots (5)$$

Formally, the distribution function of the new variates given by the relations (5) can be derived from the distribution (4), but it is much more simply obtained from the known corresponding result for normally distributed quantities.

The sample means of the normal variates are defined by

$$\bar{x}_i = \frac{1}{n} \sum_{k=1}^n x_{ik}, \quad i=1, 2, \dots, p.$$

Denoting the matrix

$$[x_{ik} - \bar{x}_i], \quad \begin{matrix} i=1, 2, \dots, p, \\ k=1, 2, \dots, n, \end{matrix}$$

by \mathbf{X} , and writing

$$\mathbf{X}\mathbf{X}' = \mathbf{C} = [C_{ij}],$$

the distribution of the means and sums of squares and products (Wishart 1928) is

$$\frac{\left| \frac{\sigma^2}{n} \mathbf{R} \right|^{-\frac{1}{2}}}{(2\pi)^{\frac{1}{2}p}} \exp \left\{ -\frac{1}{2} \bar{\mathbf{x}}' \left(\frac{\sigma^2}{n} \mathbf{R} \right)^{-1} \bar{\mathbf{x}} \right\} d\bar{\mathbf{x}} \\ \times \frac{|\sigma^2 \mathbf{R}|^{-\frac{1}{2}(n-1)} |\mathbf{C}|^{\frac{1}{2}(n-p-2)} \exp \left\{ -\frac{1}{2\sigma^2} \text{Tr} (\mathbf{R}^{-1} \mathbf{C}) \right\}}{2^{\frac{1}{2}p(n-1)} \pi^{\frac{1}{2}p(p-1)} \prod_{i=1}^p \Gamma_{\frac{1}{2}}(n-i)} \prod_{i \leq j} dC_{ij}, \quad \dots (6)$$

where $\bar{\mathbf{x}}'$ is the row vector $[\bar{x}_1 \ \bar{x}_2 \ \dots \ \bar{x}_p]$.

From (3) and (5)

$$\begin{pmatrix} \bar{t}_i = \bar{x}_i/s, \\ T_{ij} = C_{ij}/s^2, \end{pmatrix} \dots\dots\dots (7)$$

and the jacobian of (7) is $s^{p(p+2)}$. Substitute in (6), using the relations (7), multiply by the distribution of s and integrate for s from 0 to ∞ , and the distribution of the \bar{t}_i and T_{ij} takes the form

$$\frac{\left| \frac{1}{n} \mathbf{R} \right|^{-\frac{1}{2}} \Gamma_{\frac{1}{2}}(\nu + pn) \left| \mathbf{R} \right|^{-\frac{1}{2}(n-1)} \left| \mathbf{T} \right|^{\frac{1}{2}(n-p-2)}}{(\pi\nu)^{\frac{1}{2}p} \Gamma_{\frac{1}{2}}\nu \pi^{\frac{1}{2}p(p-1)} \nu^{\frac{1}{2}p(n-1)} \prod_{i=1}^p \Gamma_{\frac{1}{2}}(n-i)} \\ \times \frac{d\bar{\mathbf{t}} \prod_{i \leq j} dT_{ij}}{\left\{ 1 + \frac{\bar{\mathbf{t}}' \left(\frac{1}{n} \mathbf{R} \right)^{-1} \bar{\mathbf{t}} + \text{Tr} (\mathbf{R}^{-1} \mathbf{T})}{\nu} \right\}^{\frac{1}{2}(\nu + pn)}}, \dots\dots\dots (8)$$

where \mathbf{T} is the matrix $[T_{ij}]$. The limiting form of the distribution (8), as $\nu \rightarrow \infty$, is the Wishart distribution.

III. MARGINAL DISTRIBUTION OF THE \bar{t}_i

To obtain the marginal distribution of the means, integration of (8) with respect to the T_{ij} will be made following a method similar to that used by Cramer (1946, Section 29.5). Since \mathbf{R}^{-1} is positive definite, it may be reduced by an orthogonal matrix \mathbf{H} , so that

$$\mathbf{H} \mathbf{R}^{-1} \mathbf{H}' = \mathbf{\Lambda},$$

where $\mathbf{\Lambda}$ is diagonal, its diagonal elements λ_i being the latent roots of \mathbf{R}^{-1} . The same transformation applied to the matrix \mathbf{T} yields

$$\mathbf{H} \mathbf{T} \mathbf{H}' = \mathbf{Y}, \text{ say, } \dots\dots\dots (9)$$

transforming the $\frac{1}{2}p(p+1)$ variates in \mathbf{T} to $\frac{1}{2}p(p+1)$ new variates in \mathbf{Y} . The relation (9) represents a linear transformation of the variables whose determinant (the jacobian required) is a power of the determinant $|\mathbf{H}|$ (MacDuffee 1943; James 1954). As \mathbf{H} is orthogonal, the jacobian is unity, and, omitting the constant, (8) becomes

$$\frac{|\mathbf{Y}|^{\frac{1}{2}(n-p-2)} d\bar{\mathbf{t}} \prod_{i \leq j} dY_{ij}}{\left\{ 1 + \frac{\bar{\mathbf{t}}' \left(\frac{1}{n} \mathbf{R} \right)^{-1} \bar{\mathbf{t}} + \text{Tr} (\mathbf{\Lambda} \mathbf{Y})}{\nu} \right\}^{\frac{1}{2}(\nu + pn)}}, \dots\dots\dots (10)$$

Next, for $i \neq j$, let

$$Y_{ij} = z_{ij} \sqrt{Y_{ii} Y_{jj}}, \dots\dots\dots (11)$$

and denote the diagonal matrix, whose i th diagonal element is $\sqrt{Y_{ii}}$, by \mathbf{D} .

The jacobian of the transformation (11) is $(Y_{11} Y_{22} \dots Y_{pp})^{\frac{1}{2}(p-1)}$ and

$$|Y| = |DZD| = Y_{11} Y_{22} \dots Y_{pp} |Z|,$$

where

$$Z = \begin{vmatrix} 1 & z_{12} & \dots & z_{1p} \\ z_{21} & 1 & \dots & z_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ z_{p1} & z_{p2} & \dots & 1 \end{vmatrix}.$$

The expression (10) thus changes to

$$\frac{(Y_{11} Y_{22} \dots Y_{pp})^{\frac{1}{2}(p-3)} |Z|^{\frac{1}{2}(p-2)} d\bar{\mathbf{t}} \prod_{i=1}^p dY_{ii} \prod_{i < j} dz_{ij}}{\left\{ 1 + \frac{\bar{\mathbf{t}}' \left(\frac{1}{n} \mathbf{R} \right)^{-1} \bar{\mathbf{t}} + \text{Tr}(\Lambda Y)}{\nu} \right\}^{\frac{1}{2}(\nu + pn)}} \dots \quad (12)$$

Integration of (12) over the domain where Z is positive definite yields the factor

$$\frac{\pi^{\frac{1}{2}p(p-1)}}{\{\Gamma_{\frac{1}{2}}(n-1)\}^p} \prod_{i=1}^p \Gamma_{\frac{1}{2}}(n-i),$$

and integration over the range 0 to ∞ for each of the Y_{ii} gives

$$\frac{\nu^{\frac{1}{2}p(n-1)} \{\Gamma_{\frac{1}{2}}(n-1)\}^p \Gamma_{\frac{1}{2}}(\nu+p)}{|\Lambda|^{\frac{1}{2}(n-1)} \Gamma_{\frac{1}{2}}(\nu+pn)},$$

leaving the marginal distribution of the \bar{t}_i in the form

$$\frac{\left| \frac{1}{n} \mathbf{R} \right|^{-\frac{1}{2}} \Gamma_{\frac{1}{2}}(\nu+p)}{(\pi\nu)^{\frac{1}{2}p} \Gamma_{\frac{1}{2}}\nu} \cdot \frac{d\bar{\mathbf{t}}}{\left\{ 1 + \bar{\mathbf{t}}' \left(\frac{1}{n} \mathbf{R} \right)^{-1} \bar{\mathbf{t}} / \nu \right\}^{\frac{1}{2}(\nu+p)}}, \dots \quad (13)$$

which is the multivariate t -distribution, characterized by the matrix $\{(1/n)\mathbf{R}\}^{-1}$. The variance-covariance matrix of the \bar{t}_i is thus $(1/n)\{\nu/(\nu-2)\}\mathbf{R}$, and the marginal distribution of any variate \bar{t}_i is

$$\frac{n^{\frac{1}{2}} \Gamma_{\frac{1}{2}}(\nu+1)}{(\pi\nu)^{\frac{1}{2}} \Gamma_{\frac{1}{2}}\nu} (1 + n\bar{t}_i^2/\nu)^{-\frac{1}{2}(\nu+1)} d\bar{t}_i$$

(Cornish loc. cit.). The variate $\bar{t}_i \sqrt{n}$ is then distributed in Student's distribution with ν degrees of freedom.

IV. MARGINAL DISTRIBUTION OF THE T_{ij}

Since $\left(\frac{1}{n}\mathbf{R}\right)^{-1}$ is a positive definite symmetric matrix, the distribution (8) may be integrated easily over the range $-\infty$ to ∞ for each of the variables in $\bar{\mathbf{t}}$, yielding the numerical factor

$$\frac{\left|\frac{1}{n}\mathbf{R}\right|^{\frac{1}{2}}(\pi\nu)^{\frac{1}{2}p}\Gamma_{\frac{1}{2}}^{\frac{1}{2}}[\nu+p(n-1)]}{\Gamma_{\frac{1}{2}}^{\frac{1}{2}}(\nu+pn)}.$$

and hence the marginal distribution of the T_{ij} is

$$\frac{|\mathbf{R}|^{-\frac{1}{2}(n-1)}\Gamma_{\frac{1}{2}}^{\frac{1}{2}}[\nu+p(n-1)]}{\pi^{\frac{1}{2}p(p-1)}\nu^{\frac{1}{2}p(n-1)}\Gamma_{\frac{1}{2}}^{\frac{1}{2}}\nu\prod_{i=1}^p\Gamma_{\frac{1}{2}}^{\frac{1}{2}}(n-i)} \cdot \frac{|\mathbf{T}|^{\frac{1}{2}(n-p-2)}\prod_{i\leq j}dT_{ij}}{\{1+\text{Tr}(\mathbf{R}^{-1}\mathbf{T})/\nu\}^{\frac{1}{2}[\nu+p(n-1)]}} \quad \dots (14)$$

The substitutions $(n-1)s_i^2=T_{ii}$ and $(n-1)w_{ij}=T_{ij}$ in (14) then give the simultaneous marginal distribution of the variances and covariances.

From (14) the marginal distribution of T_{ii} is

$$\frac{\Gamma_{\frac{1}{2}}^{\frac{1}{2}}(\nu+n-1)}{\nu^{\frac{1}{2}(n-1)}\Gamma_{\frac{1}{2}}^{\frac{1}{2}}\nu\Gamma_{\frac{1}{2}}^{\frac{1}{2}}(n-1)} \cdot \frac{T_{ii}^{\frac{1}{2}(n-3)}dT_{ii}}{(1+T_{ii}/\nu)^{\frac{1}{2}(\nu+n-1)}}, \quad \dots (15)$$

which is equivalent to the distribution of Fisher's z with $(n-1)$ and ν degrees of freedom. The χ^2 distribution, as the distribution of the sum of the squares of independent normal variates in standard measure, is thus replaced by the distribution of z as the sum of the squares of uncorrelated but dependent t -variates.

Writing $T_{ii}=(n-1)s_i^2$ in (15) will give the marginal distribution of the variance, and from this the distribution of the average conditional variance of any admissible order (partial variance in Bartlett's (1933) terminology) can be obtained by appropriately adjusting the degrees of freedom involving the sample number n .

V. MARGINAL DISTRIBUTIONS OF CORRELATION AND REGRESSION COEFFICIENTS

Since regression and correlation coefficients, both total and partial, and the multiple correlation coefficient are essentially ratios of quantities involving the t -variates, in which the standardizing variate, the estimated standard deviation, cancels out, their distributions are identical with those of their counterparts relating to normally distributed quantities. These distributions may, however, be obtained directly, and as an illustration we give the derivation of the distribution of the regression coefficient, from the marginal distribution (14) when $p=2$.

With this value of p , (14) reduces to

$$\frac{\Gamma_{\frac{1}{2}}^1[\nu+2(n-1)]}{\pi^{\frac{1}{2}}\nu^{(n-1)}(1-\rho^2)^{\frac{1}{2}(n-1)}\Gamma_{\frac{1}{2}}^1\nu\Gamma_{\frac{1}{2}}^1(n-1)\Gamma_{\frac{1}{2}}^1(n-2)} \\ \times \frac{\left| \begin{matrix} T_{11} & T_{12} \\ T_{12} & T_{22} \end{matrix} \right|^{\frac{1}{2}(n-4)} dT_{11}dT_{12}dT_{22}}{\left\{ 1 + \frac{1}{\nu(1-\rho^2)}(T_{11}-2\rho T_{12}+T_{22}) \right\}^{\frac{1}{2}[\nu+2(n-1)]}} \dots \quad (16)$$

Substitute

$$T_{12}=b_{21}T_{11}$$

in (16), and take the integral with respect to T_{11} and T_{22} . Hence

$$\iint \frac{T_{11}^{\frac{1}{2}(n-2)}(T_{22}-b_{21}^2T_{11})^{\frac{1}{2}(n-4)}dT_{11}dT_{22}db_{21}}{\left[1 + \frac{1}{\nu(1-\rho^2)}\{T_{22}+T_{11}(1-2\rho b_{21})\} \right]^{\frac{1}{2}[\nu+2(n-1)]}}$$

has to be taken over the region defined by the inequality

$$T_{22}-b_{21}^2T_{11} \geq 0.$$

Change the variable to $x=T_{22}-b_{21}^2T_{11}$, and integrate with respect to both x and T_{11} from 0 to ∞ , and the distribution of b_{21} takes the form

$$\frac{(1-\rho^2)^{\frac{1}{2}(n-1)}\Gamma_{\frac{1}{2}}^1n}{\pi^{\frac{1}{2}}\Gamma_{\frac{1}{2}}^1(n-1)}(1-2\rho b_{21}+b_{21}^2)^{-\frac{1}{2}n}db_{21},$$

as first found for normally distributed variates by Pearson (1926) and Romanovsky (1926).

VI. MARGINAL DISTRIBUTION OF THE COVARIANCE

As the starting point for the derivation of the marginal distribution of the covariance, we use the result for normally distributed quantities (Pearson, Jeffery, and Elderton 1929; Wishart and Bartlett 1932). With the notation of Section II, and taking $p=2$, the distribution of C_{12} is

$$\frac{(1-\rho^2)^{\frac{1}{2}(n-3)} \exp \{ \rho C_{12} / \sigma^2 (1-\rho^2) \} \left\{ \frac{|C_{12}|}{\sigma^2 (1-\rho^2)} \right\}^{\frac{1}{2}(n-2)} K_{\frac{1}{2}(n-2)} \left\{ \frac{|C_{12}|}{\sigma^2 (1-\rho^2)} \right\} dC_{12}}{\pi^{\frac{1}{2}} \sigma^2 2^{\frac{1}{2}(n-2)} \Gamma_{\frac{1}{2}}^1(n-1)} \dots \quad (17)$$

where the vertical bars now designate the modulus, and $K_m(x)$ is the Bessel function of the second kind with imaginary argument.

For a fixed value of s^2 , make the substitution $C_{12}=s^2T_{12}$ in (17), multiply by the distribution of s^2 , and integrate with respect to s^2 from 0 to ∞ , obtaining

$$\frac{\nu^{\frac{1}{2}}\nu(1-\rho^2)^{\frac{1}{2}(n-3)}dT_{12}}{\pi^{\frac{1}{2}}2^{\frac{1}{2}(\nu+n-2)}\sigma^{\nu+2}\Gamma_{\frac{1}{2}}^1\nu\Gamma_{\frac{1}{2}}^1(n-1)}\int_0^\infty (s^2)^{\frac{1}{2}\nu} \exp \left\{ -\frac{s^2}{2\sigma^2} \left(\nu - \frac{2\rho T_{12}}{1-\rho^2} \right) \right\} \\ \times \left\{ \frac{s^2 |T_{12}|}{\sigma^2 (1-\rho^2)} \right\}^{\frac{1}{2}(n-2)} K_{\frac{1}{2}(n-2)} \left\{ \frac{s^2 |T_{12}|}{\sigma^2 (1-\rho^2)} \right\} d(s^2). \dots \quad (18)$$

A second substitution

$$y = \frac{s^2 |T_{12}|}{\sigma^2(1-\rho^2)}$$

reduces the integral in this expression to

$$\int_0^\infty y^{\frac{1}{2}(\nu+n-2)} \exp \left\{ -\frac{1-\rho^2}{2|T_{12}|} \left(\nu - \frac{2\rho T_{12}}{1-\rho^2} \right) y \right\} K_{\frac{1}{2}(n-2)}(y) dy,$$

and this may be evaluated using the result given by Watson (1922, p. 388 (7)), yielding

$$\frac{(\frac{1}{2}\pi)^{\frac{1}{2}} \Gamma_{\frac{1}{2}}(\nu+2) \Gamma_{\frac{1}{2}}[\nu+2(n-1)] P_{\frac{1}{2}(n-3)}^{-\frac{1}{2}(\nu+n-1)}(\cosh \alpha)}{(\sinh \alpha)^{\frac{1}{2}(\nu+n-1)}}, \quad \dots \quad (19)$$

where $\cosh \alpha = \nu(1-\rho^2) - 2\rho T_{12}/(2|T_{12}|)$ and $P_{\mu}^{\lambda}(x)$ is the associated Legendre function of the first kind.

Multiplying (19) by the constant from (18) and substituting for the hyperbolic functions reduces the distribution of T_{12} to the form

$$\frac{\nu^{\frac{1}{2}(\nu+2)} \Gamma_{\frac{1}{2}}[\nu+2(n-1)]}{2 \Gamma_{\frac{1}{2}}(n-1)} \cdot \frac{(1-\rho^2)^{\frac{1}{2}(\nu+n-1)} |T_{12}|^{\frac{1}{2}(n-3)}}{[\{\nu(1-\rho^2) - 2\rho T_{12}\}^2 - 4|T_{12}|^2]^{\frac{1}{2}(\nu+n-1)}} \\ \times P_{\frac{1}{2}(n-3)}^{-\frac{1}{2}(\nu+n-1)} \left\{ \frac{\nu(1-\rho^2) - 2\rho T_{12}}{2|T_{12}|} \right\} dT_{12}. \quad \dots \dots \dots (20)$$

The change of variable $T_{12} = (n-1)w_{12}$ will give the marginal distribution of the covariance, and from this the distribution of the partial covariance of any admissible order can be obtained by appropriately adjusting the degrees of freedom involving the sample number n .

VII. REFERENCES

- BARTLETT, M. S. (1933).—*Proc. Roy. Soc. Edinb.* **53** (3): 260.
 BECHHOFFER, R. E., DUNNETT, C. W., and SOBEL, M. (1954).—*Biometrika* **41**: 170.
 CORNISH, E. A. (1954).—*Aust. J. Phys.* **7**: 531.
 CRAMER, H. (1946).—"Mathematical Methods of Statistics." (Princeton Univ. Press.)
 DUNNETT, C. W., and SOBEL, M. (1954).—*Biometrika* **41**: 153.
 JAMES, A. T. (1954).—*Ann. Math. Statist.* **25**: 40.
 MACDUFFEE, C. C. (1943).—"Vectors and Matrices." (Math. Assoc. Amer.)
 PEARSON, K. (1926).—*Proc. Roy. Soc. A* **122**: 1.
 PEARSON, K., JEFFERY, G. B., and ELDERTON, E. M. (1929).—*Biometrika* **21**: 164.
 ROMANOVSKY, V. (1926).—*Bull. Acad. Sci. U.R.S.S.* [6] **10**: 643.
 WATSON, G. N. (1922).—"Theory of Bessel Functions." (Cambridge Univ. Press.)
 WISHART, J. (1928).—*Biometrika* **20A**: 32.
 WISHART, J., and BARTLETT, M. S. (1932).—*Proc. Camb. Phil. Soc.* **28**: 455.

CHORD CONSTRUCTION FOR CORRECTING AERIAL SMOOTHING

By R. N. BRACEWELL*

[Manuscript received January 28, 1955]

Summary

A graphical method is given for correcting the smoothing effect which arises when a celestial distribution of brightness temperature is scanned with an aerial beam. The method is much simpler to apply than previous methods which have been used and is no less accurate.

I. INTRODUCTION

The general problem of aerial smoothing in radio astronomy has been discussed by Bracewell and Roberts (1954) with a view to clarifying the situation. In that paper current practice as regards the correction of smoothing was critically examined. In the present communication a new method of correction is proposed which has the merits of being quick to apply and accurate.

II. GRAPHICAL METHOD OF CORRECTION

The correction is carried out graphically as follows.

In Figure 1 let $T_a(\varphi)$ be an observed temperature distribution which it is desired to correct. Let $T(\varphi)$ be the true distribution and $A(\varphi)$ the response of the aerial to a point source. Then the integral equation connecting these functions is

$$T(\varphi) = \int_{-\infty}^{\infty} A(\varphi - u)T(u)du.$$

It has been shown (Bracewell 1955) that a solution of this equation, given $T_a(\varphi)$ and $A(\varphi)$, can be expressed in the form

$$T_a(\varphi) + \sum \psi_n \Delta^n T_a(\varphi),$$

where Δ^n signifies the operation of taking n th-order finite differences over an interval a . When $A(\varphi)$ is symmetrical about $\varphi=0$, n ranges over the positive even integers. If only two terms of the solution are retained we have

$$T_a(\varphi) + \psi_2 \Delta^2 T_a(\varphi). \quad \dots\dots\dots (1)$$

This expression may be readily evaluated graphically, provided ψ_2 and a are stated. Thus Figure 1 shows a chord of span $2a$ drawn across an observed curve T_a . The mid point of the chord falls below the curve by an amount proportional to $\Delta^2 T_a(\varphi)$. Hence, provided ψ_2 is known, the value $T_a + \psi_2 \Delta^2 T_a$

* Fulbright Visiting Professor in the Astronomical Department, University of California, Berkeley; on leave from Radiophysics Laboratory, C.S.I.R.O., Sydney.

may be plotted. This is a particularly simple construction, especially as, in general, it is possible to arrange that $\psi_2 = -\frac{1}{2}$. Then the amount of the correction is exactly equal to the intercept between the chord and the curve, thus eliminating any multiplication from the construction.

The necessary quantities ψ_2 and a have to be determined from $A(\varphi)$, and prove to be independent of φ . It has been shown (Bracewell 1955) that in cases arising in optics one may take $\psi_2 = -\frac{1}{2}$ and $a = \sigma$ where σ is the standard deviation of the instrumental profile. This possibility vanishes when the instrumental profile is an aerial pattern, for the following reason. The aerial aperture distribution, being necessarily finite, in general falls discontinuously to zero at the

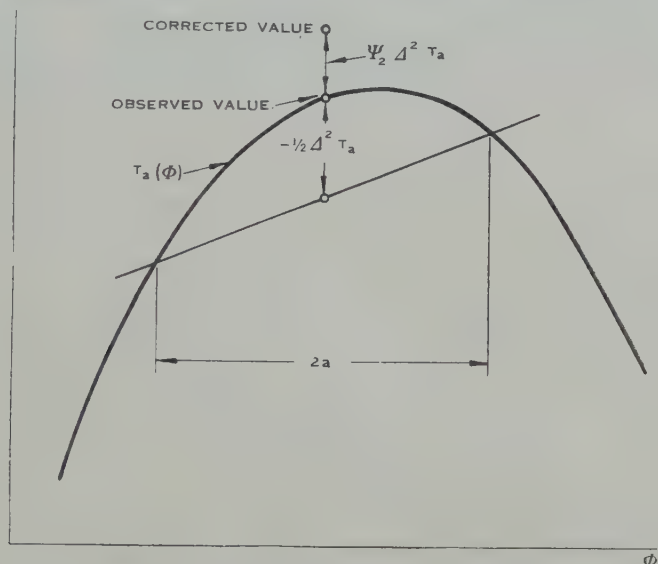


Fig. 1.—Details of the chord construction.

extremes of the aerial. The Fourier transform of such a distribution will behave asymptotically as φ^{-1} when $\varphi \rightarrow \infty$. The aerial pattern $A(\varphi)$, which is the square of this transform, will behave as φ^{-2} . Consequently the second moment of the distribution $A(\varphi)$ about $\varphi = 0$ will not, in general, exist. This circumstance necessitates a special approach to the case of aerial smoothing as distinct from other types of instrumental broadening.

By the convolution theorem,

$$\bar{T}_a(s) = \bar{A}(s) \bar{T}(s),$$

where the bars indicate Fourier transforms, and therefore

$$\bar{T}(s) = \bar{T}_a(s) + \left(\frac{1}{\bar{A}(s)} - 1 \right) \bar{T}_a(s) \quad \dots \dots \dots (2)$$

for values of s such that $A(s) \neq 0$. Now if we compare the right-hand side of (2) with the Fourier transform of (1), namely,

$$\bar{T}_a(s) - (4\psi_2 \sin^2 \pi a s) \bar{T}_a(s),$$

it will be seen that the determination of ψ_2 and a reduces to arranging the best fit between $4\psi_2 \sin^2 \pi as$ and $[\bar{A}(s)]^{-1} - 1$, and this may be done graphically as in the examples that follow.

(a) *Example 1: Uniform Aperture Distribution*

In this common case we have

$$A(\varphi) = b \left(\frac{\sin \pi \varphi / b}{\pi \varphi} \right)^2,$$

and therefore

$$\bar{A}(s) = 1 - bs, \quad 0 < s < b^{-1}. \quad \dots \dots \dots (3)$$

In Figure 2 we see $[\bar{A}(s)]^{-1} - 1$ and $2 \sin^2 \frac{1}{2} \pi bs$ compared. The agreement is good for values of bs as large as $\frac{1}{2}$. Hence excellent correction may be expected

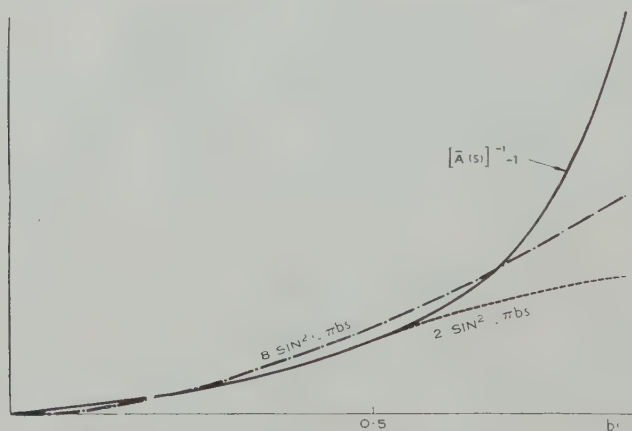


Fig. 2.—Showing the good agreement between $[\bar{A}(s)]^{-1} - 1$ and $2 \sin^2 \frac{1}{2} \pi bs$.

for Fourier components such that $0 < s < \frac{1}{2} b^{-1}$, where b^{-1} is the cut-off value of s (see Bracewell and Roberts 1954), and partial correction may be expected for components of higher spatial frequency up to b^{-1} . As shown by Bracewell and Roberts $T_a(\varphi)$ will not contain Fourier components with more than b^{-1} waves per unit of φ .

For this case, then, we find

$$\begin{aligned} \psi_2 &= -\frac{1}{2}, \\ a &= \frac{1}{2}b. \end{aligned}$$

It is useful to note that the chord span $2a$ is equal to half the beam width between zeros.

Other ways of choosing ψ_2 and a are open, and in special circumstances they may have value. For example, the curve $8 \sin^2 \frac{1}{4} \pi bs$ is shown by a dot-dash line in Figure 2. In this case the chord span is one-half the former value but the factor ψ_2 is four times larger. Probably this would not be as convenient in practice in most cases.

(b) *Example 2: Tapered Distribution*

The distribution in the aperture plane of a paraboloidal reflector normally falls off towards the rim and we may take for an example the aperture distribution of Figure 3, namely,

$$E(x) = 2 - x^2, \quad |x| < 1.$$

The function $\bar{A}(s)$, which is proportional to the autocorrelation function of $E(x)$, is readily calculated. Although in the present case the integral can be performed without difficulty, it will be found much quicker to calculate $\bar{A}(s)$ numerically (Fig. 3 (b)). The resulting curve of $[\bar{A}(s)]^{-1} - 1$ is compared with the uniform case in Figure 3 (c). The agreement is seen to be rather close, and further examination shows that $2 \sin^2 \frac{1}{2} \pi bs$ is as good a match for the tapered as for the uniform case.

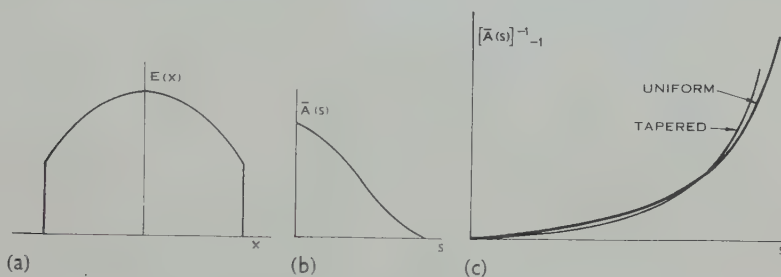


Fig. 3 (a).—Tapered aperture distribution.

Fig. 3 (b).—Corresponding $\bar{A}(s)$.

Fig. 3 (c). $[\bar{A}(s)]^{-1} - 1$ for the uniform and tapered aperture distributions.

It therefore appears that moderate tapering of the aperture distribution does not seriously affect ψ_2 and a . However, in a particular case this matter could always be studied very easily by the method indicated, with a view to modifying ψ_2 and a to obtain a best fit over the range of s considered most appropriate to the nature of $T_a(\varphi)$.

III. COMPARISON WITH THE METHOD OF SUCCESSIVE SUBSTITUTIONS

The method of correcting aerial smoothing which has been most widely adopted in radio astronomy (Bolton and Westfold 1950; Bracewell and Roberts 1954) is referred to as the method of successive substitutions. In this method one tests an approximate solution by subjecting it to aerial smoothing and compares the result with $T_a(\varphi)$. The discrepancy is then applied as a correction to the approximate solution, and the process repeated. This method engenders confidence, because it is self-checking, and is quite popular in fields other than radio astronomy (van Cittert 1931). One or two stages are usually found to bring the discrepancy within the observational error.

In the uniform aperture case, the spectrum of T_a bears a ratio to that of T which is the function of s shown in curve A of Figure 4 (see equation (3)). After one stage of correction this spectral factor is brought closer to unity (curve B) and after two and three stages the effect is as represented by the dotted curves.

Curve D represents the effect of a very large number of stages, i.e. full restoration. The equation of the n -stage curve may be shown to be

$$y = 1 - (bs)^{n+1}.$$

With the chord construction here described, the degree of restoration jumps straight to the high values of curve C whose equation is

$$y = (1 - bs)(1 + 2 \sin^2 \frac{1}{2} \pi bs).$$

This impressive result shows that the simple chord construction is as good as several stages of successive substitutions. Evidently the series of finite differences converges more rapidly than the series on which the method of successive substitutions is based.

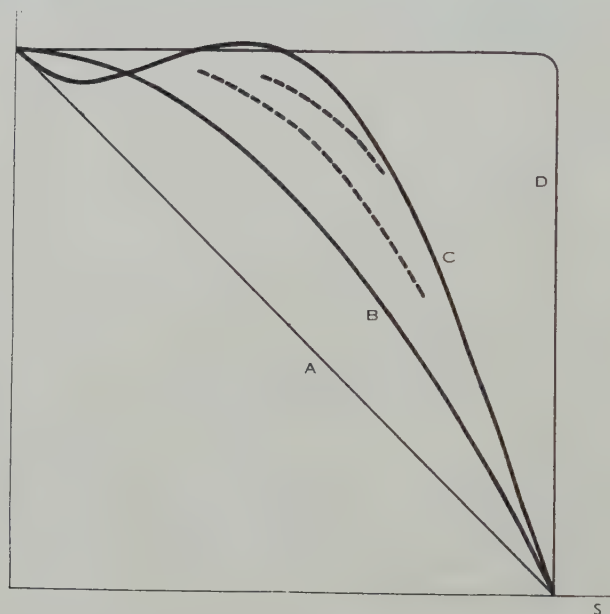


Fig. 4.—Spectral content of various distributions relative to that of $T(\varphi)$.

It should not be forgotten, however, that the precise nature of $T_a(\varphi)$ also enters into a comparison between methods. Thus for small values of s it will be noticed that, although the corrections are small, nevertheless successive substitutions give them better than the chord method. Therefore, to take an extreme case, if $T_a(\varphi)$ contains strong components for $s < 0.2$, and weak components for $s > 0.2$, then better results might be obtained by successive substitutions.

As a comparative illustration we may take the stripwise distribution across the quiet Sun at 21 cm observed by Christiansen and Warburton (1953). Figure 5 shows $T_a(\varphi)$, $A(\varphi)$, and the chord used for correction. Curve (a), shown displaced, is the result of correcting by the chord construction, and curve (b) shows for comparison the known result of full restoration calculated by Bracewell and

Roberts. The present result is, for most purposes, as good as the full restoration, and better than the result of one stage of successive substitutions. In this example the low values of s are present in considerable strength since the distribution is wide compared with the aerial beam width. The effect discussed in the previous paragraph therefore appears and its magnitude will be looked into. The effect is to give zero correction over the flat top of $T_a(\phi)$, and this leads to values about 2 per cent. too low in this range of ϕ . Thus the effect is

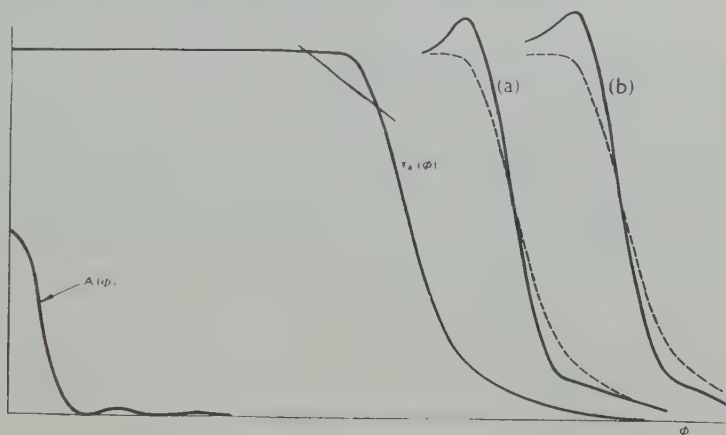


Fig. 5.—Stripwise distribution of brightness across the quiet Sun at 21 cm; (a) corrected by chord construction, (b) accurate solution.

unimportant, being within the limits of error. Furthermore, it has no effect on the *shape* of the distribution. It therefore seems likely that experience with the chord construction will show it to be very generally satisfactory in accuracy as well as in speed.

IV. REFERENCES

- BOLTON, J. G., and WESTFOLD, K. C. (1950).—*Aust. J. Sci. Res. A* **3**: 19.
 BRACEWELL, R. N. (1955).—*Aust. J. Phys.* **8**: 54.
 BRACEWELL, R. N., and ROBERTS, J. A. (1954).—*Aust. J. Phys.* **7**: 615.
 CHRISTIANSEN, W. N., and WARBURTON, J. A. (1953).—*Aust. J. Phys.* **6**: 262.
 VAN CITTERT, P. H. (1931).—*Z. Phys.* **69**: 298.

THE TRANSPORT OF HEAT THROUGH NARROW CHANNELS BY LIQUID HELIUM II

By P. G. KLEMENS*

[*Manuscript received December 17, 1954*]

Summary

In narrow channels ($\sim 10^{-4}$ cm) the observed heat transport is considerably larger than calculated by the internal convection theory. It is suggested that, because of the anisotropy of the distribution of phonons in the channel walls resulting from the temperature gradient, the normal fluid in the immediate vicinity of the walls is not at rest but flows towards the colder region. The magnitude of the resulting heat transport is in reasonable agreement with the observed discrepancy.

I. INTRODUCTION

Heat transport in liquid helium II is considered to be an internal convection process (London and Zilsel 1948 ; Gorter and Mellink 1949). A temperature gradient gives rise to a fountain pressure gradient

$$\text{grad } p = \rho S \text{ grad } T, \quad \dots\dots\dots (1)$$

where ρ is the total density of the fluid and S its entropy per unit mass. The fountain pressure drives the normal fluid towards the cold region, and there is an equal return flow of superfluid. In narrow channels and for small temperature gradients this flow is limited by the viscous friction of the normal fluid. We shall not consider here cases of wide channels and large temperature gradients, where the mutual friction between the normal fluid and the superfluid becomes important, leading to a non-linear dependence of flow on $\text{grad } T$ (Gorter and Mellink 1949). The flow (volume per unit time) through a slit of unit breadth and width d due to the fountain pressure is

$$V = -\frac{d^3}{12\eta} \text{ grad } p, \quad \dots\dots\dots (2)$$

where η is the coefficient of viscosity of the normal fluid. Since the returning superfluid has no entropy, this circulation mechanism gives rise to a heat current

$$Q = \rho S T V = -\frac{\rho^2 S^2 T d^3}{12\eta} \text{ grad } T \quad \dots\dots\dots (3)$$

leading to an apparent thermal conductivity $\kappa = -Q/(d \text{ grad } T)$, which varies with temperature and is proportional to d^2 .

When comparing the heat flow measured by Keesom and Duyckaerts (1947) and by Meyer and Mellink (1947) against (3), as was done by London

* Division of Physics, C.S.I.R.O., University Grounds, Sydney.

and Zilsel (1948), it is found that the observed heat flow exceeds the values calculated from (3) for slit widths of order 10^{-4} cm or less, the relative discrepancy becoming larger for small slit widths and low temperatures (see Table 1, column 5).

Gorter and Mellink (1949) have suggested that this discrepancy arises because the slit width d is less than the mean free path l_0 of those excitations in the fluid which determine the bulk viscosity, and the heat flow is increased by a factor of order l_0/d . However, the observed thermal conductivities do not seem to vary as d , as required by this explanation.

An alternative explanation of this discrepancy is suggested here. In the derivation of (2) the assumption has been made that the normal fluid in the immediate vicinity of the channel walls is at rest. This will only be so if the walls are isothermal, for, if the normal fluid exchanges momentum with the phonon gas at the walls, an anisotropic phonon distribution will tend to induce in the normal fluid a bulk flow tangential to the walls in the opposite direction to the temperature gradient which causes the phonon anisotropy. This additional flow is again compensated by a return flow of superfluid, and the additional heat transport can account for the discrepancies between the observed flow and equation (3).

II. QUASI-EQUILIBRIUM BETWEEN THE FLUID AND THE PHONONS OF THE WALL

It can be shown (Klemens 1951) that the distribution of phonons in a temperature gradient is of the form

$$N_{\mathbf{k}} = [e^{\alpha\omega - \lambda \cdot \mathbf{k}} - 1]^{-1} \simeq [e^{\alpha\omega} - 1]^{-1} + \lambda \cdot \mathbf{k} e^{\alpha\omega} [e^{\alpha\omega} - 1]^{-2}, \quad \dots (4)$$

where N is the average occupation of the normal mode of frequency ω and wave-vector \mathbf{k} , and $\alpha = \hbar/2\pi KT$. The vector λ characterizes the anisotropy of the distribution, and is given by

$$\lambda = -\text{grad } T \frac{l}{T} \frac{\hbar c}{2\pi KT}, \quad \dots (5)$$

where l is the phonon mean free path and c the velocity of sound.

The distribution function of the bosons which describe the fluid is of the form

$$N_{\mathbf{k}} = [e^{(E - E_0)/KT - \lambda' \cdot \mathbf{k}} - 1]^{-1}, \quad \dots (6)$$

where E is the energy and \mathbf{k} the wave-vector of the boson state considered, and E_0 is a parameter. If the fluid is at rest relative to the observer, $\lambda' = 0$, leading to the normal Planck distribution. In general, (6) is the distribution of a fluid moving relative to the frame of reference with a bulk velocity \mathbf{v} given by

$$\lambda' = \hbar \mathbf{v} / 2\pi KT \quad \dots (7)$$

(e.g. Dingle 1952, p. 137).

If in a system of interacting bosons energy is conserved for each interaction process, it is easily seen that equation (6), with $\lambda' = 0$, describes a stationary

state; if two subsystems interact so as to conserve energy, they will tend to a mutual equilibrium with a common value of T in their distribution functions. Similarly, if both energy and momentum (that is, \mathbf{k}) are conserved in individual interactions, (6) will be a stationary distribution for any set of parameters T and λ' , and two subsystems, conserving energy and momentum in each interaction process, will tend to a quasi-equilibrium with common values of T and λ' in their distribution functions.

If the interactions of the bosons of the fluid with the phonons of the wall conserve the tangential component of \mathbf{k} , then these will be in mutual equilibrium if $\lambda = \lambda'$ in their distribution functions (6) and (4) and if the component of λ normal to the wall vanishes. Under isothermal conditions $\lambda = 0$, hence $\lambda' = 0$ for the normal fluid in the immediate vicinity of the walls, and it is therefore at rest. But in the presence of a temperature gradient $\lambda \neq 0$, and, to satisfy $\lambda = \lambda'$, the bulk velocity of the normal fluid near the wall will, from (5) and (7), be given by

$$\mathbf{v} = -\frac{cl}{T} \text{grad } T. \quad (8)$$

With (8) as boundary condition instead of $\mathbf{v} = 0$, there will be an additional flow through the slit given by

$$V' = -d \frac{cl}{T} \text{grad } T, \quad (9)$$

and thus an additional heat flow given by

$$Q' = \rho S T V' = -cd\rho S \text{grad } T, \quad (10)$$

which corresponds to an additional thermal conductivity

$$\kappa' = S\rho cl, \quad (11)$$

which is independent of the cross section of the channel.

III. COMPARISON WITH OBSERVATIONS

In Figure 1 are plotted differences between the thermal conductivities observed for various slit widths (κ_0) and the conductivities calculated from (3) ($\kappa_{(3)}$). According to the above theory this difference should be independent of slit width, and a function of temperature only. Considering the uncertainties of the observations of Q and d , the observed points could well fall on a single curve, though the scatter for high temperatures is considerable. In any case the plotted points do not reveal any systematic dependence of κ' on slit width, so that an equation of the form (11) could account for the discrepancies.

To calculate the additional heat flow from (11), the velocity and the mean free path of the phonons in the walls must be known. In the experiments of Keesom and Duyckaerts (1947) and of Meyer and Mellink (1947) the slits were formed between two optically flat surfaces of glass. One can take for glass $c \sim 2 \times 10^5$ cm/sec. The mean free path for phonons in quartz glass is known from

measurements of the thermal conductivity at low temperatures (Berman 1951) and the theory of Klemens (1951). Thus, for longitudinal phonons of wave number k , $l \sim 3.6 \times 10^9 k^{-2}$ cm. The mean free path of transverse phonons is probably smaller by a factor of about 50.

Since λ is not a constant, λ' must be equated to some average value of λ . There arises a difficulty because the appropriate averaging procedure depends upon the nature of the interactions between the phonons of the wall and the

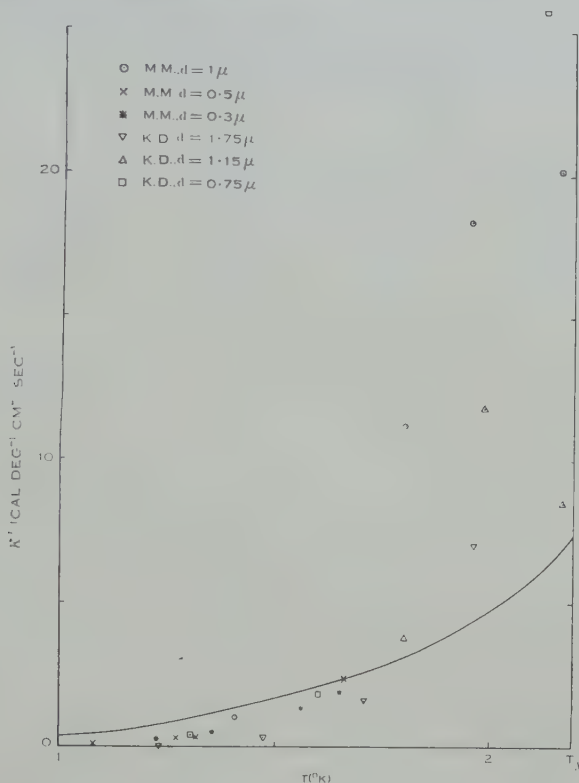


Fig. 1.—Difference between the observed conductivities and conductivities calculated from equation (3). Full line shows κ' calculated from equation (11).

bosons of the normal fluid. If one takes for this average, admittedly somewhat arbitrarily, the mean free path for the frequency $2\pi KT/h$, averaged over the three polarizations, one obtains $l = 2.8 \times 10^{-3} T^{-2}$ cm.

Values of κ' calculated from (11), using the above values of l and c , and values of S taken from Kramers, Wasscher, and Gorter (1952), are plotted as a function of temperature in Figure 1. Although the ratio $\kappa_0/(\kappa_{(3)} + \kappa')$, given in Table 1, clearly does not depart from unity to the extent that $\kappa_0/\kappa_{(3)}$ does, there is considerable lack of agreement between κ' and the experimental points. This is of such a form as to indicate that l is not as strongly temperature dependent as one would expect. This may be due to a difference between the phonon

mean free paths near a polished surface and in the interior, or it may be due to the role of the phonons of different polarizations changing with temperature. It should also be remembered that the experimental results are uncertain, particularly just below the λ -point, and that the mean free paths of phonons differ significantly from glass to glass. It is significant that α' appears to be a

TABLE 1
APPARENT HEAT CONDUCTIVITY OF HELIUM II THROUGH SLITS OF WIDTH d

α_0 values observed by Keesom and Duyckaerts (1947) (K.D.), and by Meyer and Mellink (1947)—M.M.; $\alpha(3)$ theoretical values calculated from equation (3) (London and Zilsel 1948); α' correction calculated from equation (11)

d (10^{-4} cm)	T (°K)	α_0 (cal deg $^{-1}$ cm $^{-1}$ sec $^{-1}$)	$\alpha(3)$	$\frac{\alpha_0}{\alpha(3)}$	$\frac{\alpha_0}{\alpha(3) + \alpha'}$
1.75 (K.D.)	1.960	17.2	10.2	1.69	1.18
	1.705	3.5	1.86	1.88	0.80
	1.476	0.61	0.32	1.92	0.32
	1.223	0.066	0.032	2.05	0.09
1.15 (K.D.)	2.170	24	15.5	1.55	1.18
	1.989	17.1	5.28	3.25	1.74
	1.799	5.35	1.55	3.45	1.14
1.0 (M.M.)	2.159	31	10.8	2.9	1.75
	1.948	21.4	3.06	7.0	2.9
	1.802	12.4	1.19	10.4	2.8
	1.411	1.05	0.060	17.5	0.75
0.75 (K.D.)	2.097	30.2	4.3	7.0	2.9
	1.600	1.85	0.156	11.8	0.82
	1.403	0.39	0.031	12.5	0.27
0.5 (M.M.)	1.659	2.48	0.108	23	0.99
	1.315	0.354	0.0064	55	0.35
	1.274	0.277	0.0043	64	0.31
	1.086	0.124	0.0006	202	0.28
0.3 (M.M.)	1.652	1.92	0.0365	54	0.82
	1.558	1.35	0.018	75	0.69
	1.358	0.48	0.0034	140	0.42
	1.226	0.25	0.00097	260	0.33

function of temperature independent of slit width, and that the present theory gives the observed order of magnitude of the heat flow in all cases, while equation (3) fails to do so.

It would be interesting to observe the heat flow through slits formed of materials with a different phonon mean free path, such as crystalline quartz, diamond, or artificial sapphire; the heat transport through these slits should be considerably enhanced.

IV. REFERENCES

- BERMAN, R. (1951).—*Proc. Roy. Soc. A* **208**: 90.
DINGLE, R. B. (1952).—*Advanc. Phys.* **1**: 111.
GORTER, C. J., and MELLINK, J. H. (1949).—*Physica, 's Grav.* **15**: 285.
KEESOM, W. H., and DUYCKAERTS, G. (1947).—*Physica, 's Grav.* **13**: 153.
KLEMENS, P. G. (1951).—*Proc. Roy. Soc. A* **208**: 108.
KRAMERS, H. C., WASSCHER, J. D., and GORTER, C. J. (1952).—*Physica, 's Grav.* **18**: 329.
LONDON, F., and ZILSEL, P. R. (1948).—*Phys. Rev.* **74**: 1148.
MEYER, L., and MELLINK, J. H. (1947).—*Physica, 's Grav.* **13**: 197.

A DIFFERENTIAL REFRACTOMETER OF HIGH SENSITIVITY

By P. G. GUEST*

[*Manuscript received November 10, 1954*]

Summary

The refractometer is of the Jamin type. The positions of the main and fiduciary fringes are determined by means of a pair of photomultiplier tubes, the outputs of which are fed to a balanced amplifier. A change in optical path of $1/15,000$ th of a fringe can be detected, and the instrument is stable to within $1/5000$ th of a fringe over 30 min periods. Examples are given of its use in the measurement of small differences in refractive index for liquids and gases.

I. INTRODUCTION

The invention of the photomultiplier tube has greatly extended the range of light intensities which can be detected with the aid of relatively simple equipment. In particular the intensity in the various regions of the fringe systems produced by interference methods can be readily measured. In the present paper a description will be given of a highly sensitive interference refractometer using these tubes as detectors.

The Jamin refractometer appears to be the most suitable type for this purpose because a relatively large source can be used and the fringe spacing can be easily varied. With the simpler Rayleigh refractometer the source slit is very narrow, giving rather faint fringes, and the fringes are very close together. Because of changes in the intensity of the source and drifts in the positions of the fringes with time it is essential to provide a fiduciary system of fringes with which the main system can be compared. It is still quite commonly claimed that such a system is a prerogative of the Rayleigh refractometer, but in point of fact a fiduciary system is readily obtained with the Jamin refractometer (Guest and Simmons 1953). Attention should also be drawn to a refractometer employing quite different principles designed by Ingelstam (1953), in which the cell containing the fluids is imaged in phase contrast. This instrument also employs photomultiplier tubes as detectors, and an accuracy of $1/1000$ th of a fringe is obtained.

Up to the present the refractometer described here has only been used for the measurement of small differences in refractive index—mainly those due to impurities in water samples. Such measurements can be undertaken in an ordinary laboratory provided the room temperature does not fluctuate very rapidly. In measuring large differences in refractive index it would be necessary

* Physics Department, University of Sydney.

to maintain the temperature very accurately constant. The cell length must also be known very accurately. Unfortunately, because of lack of funds, it has not been possible to extend the range of the instrument so that large differences in refractive index can be measured.

The drift of the refractometer over the period necessary to complete a reading is less than $1/5000$ th of a fringe. The accuracy obtainable in the refractive index difference $\Delta\mu$ for the two substances being compared is about 1×10^{-8} for water samples, using a cell of length 1.3 cm. For gases a sensitivity of 1×10^{-9} is obtained with a cell of length 20 cm. The accuracy for the intercomparison of air and a gas such as oxygen is lower by a factor of 50 because of the difficulty in measuring and maintaining constant the pressure and the temperature.

II. DESCRIPTION OF THE REFRACTOMETER

A schematic diagram of the refractometer is given in Figure 1. The two Jamin plates J_1 and J_2 are 1 cm thick, and the spacing between the two interfering beams is therefore also about 1 cm. A small value of plate thickness and beam separation reduces the effect of temperature gradients in the refractometer. However, a slightly greater plate thickness (perhaps up to 2 cm) would permit

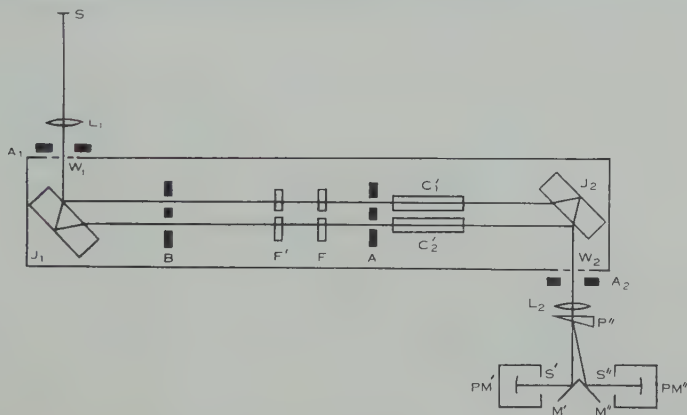


Fig. 1.—Schematic diagram of the refractometer. Components marked with a prime superscript are in the lower beam only, those with a double prime in the upper beam only.

a wider separation of the cells C_1 and C_2 , which would simplify the problem of isolating the two cells from one another. Screws are provided for tilting the plates about horizontal and vertical axes. The plates are spaced 50 cm apart to allow cells of various lengths to be used. The entire path traversed by the interfering beams is shielded by a wooden cover to eliminate air currents, the light entering and leaving by the windows W_1 and W_2 .

The source slit S , in the focal plane of the lens L_1 , is a few millimetres square. Its size is limited by the necessity for preventing the beam from spreading too much in its passage through the refractometer. The lenses L_1 and L_2 both have a focal length of 25 cm. A_1 and A_2 are vertical slits which block off multiply reflected beams. The apertures, of diameter 2 mm, in the

opaque screen A select the four beams, the lower pair passing through the cells and the upper pair being used to produce a fiduciary fringe system.

Immediately behind the lens L_2 is a prism P'' , whose angle is about 20° . This prism is only in the upper beam, and it deviates this beam on to the mirror M'' . The lower beam falls on the mirror M' . S' and S'' are horizontal slits of adjustable width in the focal plane of the lens L_2 , where the fringe systems are formed. Portions of each fringe system pass through the slits and fall on the cathodes of the photomultiplier tubes, which are enclosed in light-tight holders.

Two pairs of fringe-shifting plates, F and F' , are shown in the diagram. Both are of the Jamin type, the two plates having a constant angle of separation and rotating together on a common shaft. The plates F are in both the upper and lower beams, and are employed to test whether the two systems of fringes are coincident (see Section IV below). The plates F' are in the lower beams and hence move only the main set of fringes, leaving the fiduciary system unaltered. These plates are turned by a micrometer screw at the end of a pivoted bar, so that small rotations can be accurately measured. The pitch of the screw is 0.5 mm, and the length of the bar 30 cm. A drum fitted to the head of the screw is divided into 100 divisions, and the angle between the plates is chosen so that one division corresponds to a fringe movement of about 1/4000th of the fringe spacing.

B is a shutter which can be used to block off one or other of the pairs of beams, so that the intensities of the upper and lower beams reflected from the front of J_1 , or from the back of J_1 , can be compared. This provides a check on the stability of the equipment.

The cells C_1 and C_2 contain the fluids being compared. For liquids a stainless steel block 1.3 cm thick was ground and polished so that its surfaces were flat and parallel. Grooves were cut for the cells, and the ends were closed with optical flats which were wrung to the steel. This design is simple, and is satisfactory for the comparison of very dilute water solutions, but it is not certain that the two cells are completely isolated from one another. For gases, glass tubes 20 cm long have been used, the end windows being cemented to the glass.

A white light source may be used when the refractive index difference is very small and the dispersion is not important. However, a monochromatic source is much more suitable. A standard high pressure mercury lamp of the type used for street lighting is very convenient, the green mercury line being isolated by a monochromator and a gelatin filter. With such a high intensity source the photomultiplier voltages can be kept quite low (60–70 V per stage), and this reduces the random fluctuation in the output current. The only disadvantage of such a source is the breadth of the line, which limits the optical path difference to a small number of wavelengths. It is quite possible to use a laboratory type "Osira" lamp at somewhat higher values of photomultiplier gain (with voltages of about 90 per stage) in cases where the broadening effect might be important. With weaker sources the fringe spacing is made very wide and the slits S' and S'' are opened so that all the light in the image falls on the photomultiplier cathode.

III. THE ELECTRONIC CIRCUITS

The photomultiplier tubes are of the 931A type. Figure 2 shows the circuit for the photomultiplier power supply. The rectifier uses two 5R4GY tubes and a transformer with a 385-0-385 V secondary in what might be described as a "voltage one-and-a-half" circuit giving 1500 V D.C. output. This circuit was used because a commercial radio transformer of this type was readily available, but a simple circuit using a 1000 V transformer would be equally satisfactory. This D.C. voltage is applied to a conventional electronic regulating circuit. The screen voltage of the 6J7G tube is obtained from a divider across the unregulated side of the supply, and by adjustment of this divider the regulation may be made perfect for changes in the mains input voltage.

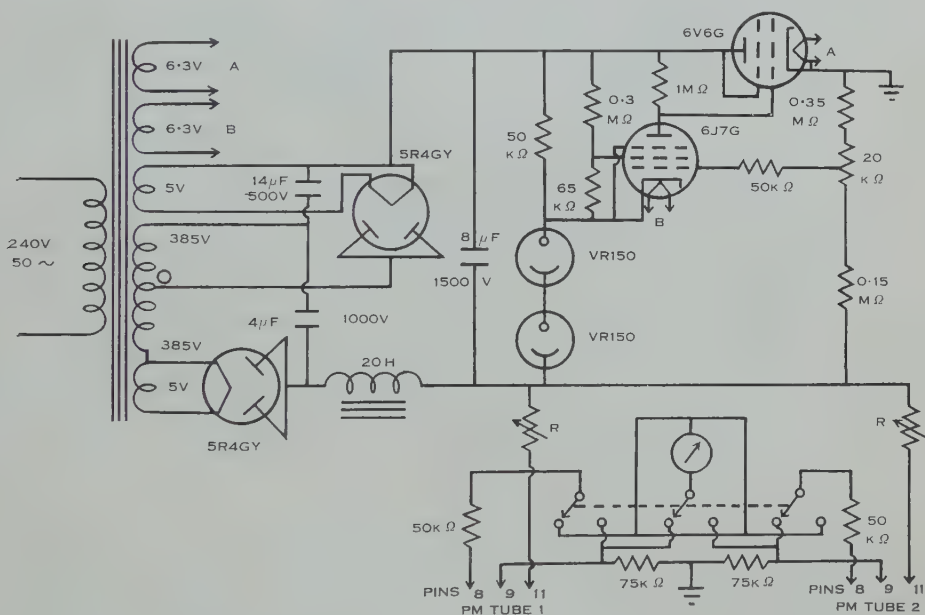


Fig. 2.—Photomultiplier power supply.

The positive side of the regulated output is earthed. The voltages for the dynodes are obtained from a series of 50,000 Ω resistors between pins 11 (cathode) and 9 (final dynode). These resistors are mounted in the 931A tube holders and are not shown in the diagram. The resistance R between pin 11 and the negative side of the supply consists of a switch for inserting resistances and a variable 50,000 Ω wire-wound rheostat. These provide the coarse and fine controls for adjusting the "volts per stage" for the 931A tubes.

The current from the network supplying the dynode voltages is applied through a switch to a calibrated meter on which the volts per stage for either tube may be read. The network is completed by a 75,000 Ω resistance to ground, and the final anode is returned to ground through the amplifier load resistance. This connection is by means of a separate shielded lead clipped

directly on to the pin, the corresponding lug on the tube socket being removed to minimize leakage currents.

The D.C. amplifier shown in Figure 3 is based on a circuit described by Scroggie (1952). As used here it is completely symmetrical. The photomultiplier currents are fed through $10\text{ M}\Omega$ load resistances connected to the grids of the EF37A tubes, and they return to earth through the small resistances in the EF55 cathode circuits. The two EF37A tubes have a common cathode resistance and the circuit behaves as a balanced amplifier. The output meter is connected between the cathodes of the two EF55 cathode followers. The EF37A may be replaced by the more modern EF86, while the two EF55 tubes may be replaced by a single 12AT7 or ECC81 double triode.

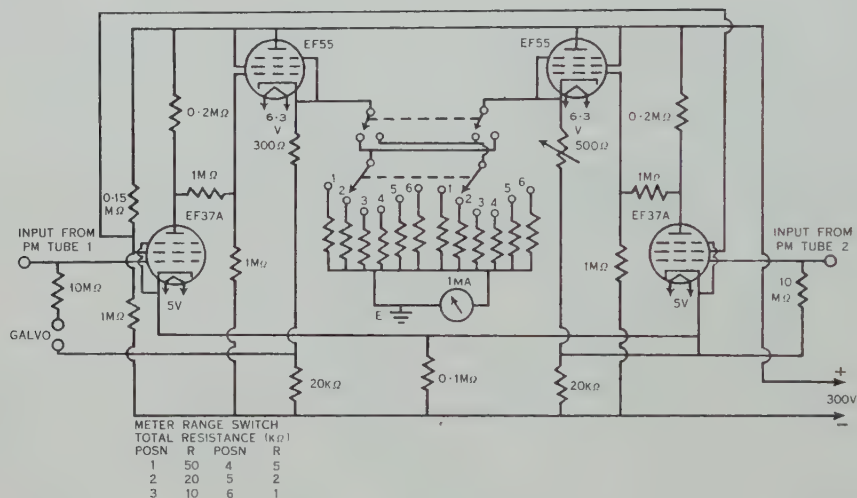


Fig. 3.—Balanced D.C. amplifier.

The range switch inserts resistances in series with the meter, which has a 1 mA movement. The centre point of the meter circuit is earthed. The meter zero position is adjusted by a $500\ \Omega$ rheostat in the cathode of one EF55 tube. The output current is a measure of the difference between the currents in the two photomultipliers. Terminals are provided for a galvanometer (sensitivity about $2\text{ cm}/\mu\text{A}$) to measure the current flowing in one of the photomultipliers.

The electronic equipment has been in operation for a period of more than two years. During that time the only trouble encountered was a change in the zero setting of the balanced amplifier, due to aging of the $1\text{ M}\Omega$ resistors coupling the plate of the EF37A tube to the grid of the EF55. These resistors should preferably be of the high stability type.

IV. ADJUSTMENT OF FRINGES TO COINCIDENCE

The galvanometer is used to measure the current produced by the fiduciary fringe system. The gain of the photomultiplier tube is adjusted till the current at the fringe maximum is $2.5\ \mu\text{A}$.

In Figure 4 are drawn typical curves of current against fringe position for the two systems as the fringes are moved across the photomultiplier tubes by the fringe-shifting plates F . In Figure 4 (b) the minima are coincident and the differences in amplitude—and hence the currents in the balanced amplifier—at positions such as those labelled 1 and 2 are the same. In Figure 4 (a) the minima are slightly displaced and the amplifier current is greater at position 1 than at position 2. In Figure 4 (c) the minima are displaced slightly in the opposite sense and the amplifier current is greater at position 2. This then provides a method for detecting when the main system of fringes is coincident with the fiduciary system, the main system being moved by means of the fringe-shifting plates F' until the amplifier currents in positions 1 and 2 are the same.

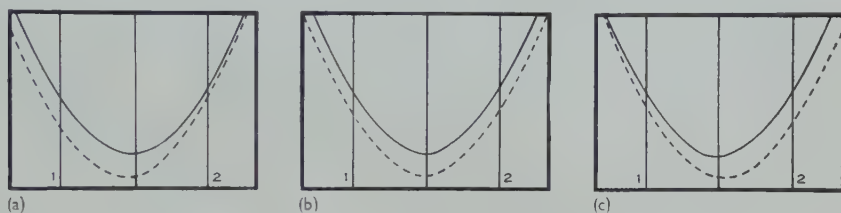


Fig. 4.—Curves of photomultiplier current against fringe position. Solid curves refer to the fiduciary system, dotted curves to the main system. In diagram (b) the minima are coincident, in the other two diagrams they are displaced.

The test positions 1 and 2 are located by moving the plates F' until the galvanometer current reaches a definite value—say $1\ \mu\text{A}$. Clearly the magnitude of the main fringe current will not affect the equality at positions 1 and 2 in Figure 4 (b). However, it is an advantage to have the current difference at these positions close to that at the minimum, for then the current difference will not vary greatly as the plates F are rotated and it will not be necessary to set the galvanometer current accurately to $1\ \mu\text{A}$.

The procedure for determining the setting of F' for which the fringe systems are coincident is then as follows. The fringe systems are moved by F' from a position corresponding to a fiduciary current of $1\ \mu\text{A}$ through the minimum till the current is again $1\ \mu\text{A}$. The difference between the balanced amplifier currents at these two positions is noted. This difference is reduced to zero by rotating F' . When the fringes are coincident the gain of the photomultiplier tube in the main fringe system may be varied till the current at the minimum is the same as at the two test positions 1 and 2. The adjustment of the gain is not at all critical and need only be performed occasionally.

Because of screw backlash it is not easy to set the fringe-shifting plates F' exactly on the coincident position. It has been found best to move the screw through intervals of one division at a time, observing the amplifier current difference for positions 1 and 2 and stopping when this difference changes sign. The correct setting can then be found by interpolation.

This method of setting to coincidence is not affected by absorption in the fluids. A second method of measurement has been used occasionally when the difference in absorption of the standard and unknown fluids is very small.

The fringes are set in some position such as 2 in Figure 4 (*b*) and the gains and the plates F' adjusted till a movement of the fringe systems about this position leaves the amplifier current unaltered. The two curves have then the same slope in the region about position 2, and the amplifier current is unaffected by simultaneous drifts of the two sets of fringes. It is, however, affected by changes in the photomultiplier sensitivities and by zero drift in the amplifier, and these must be allowed for. If the pair of beams reflected from the front of the plate J_1 is blocked off by the shutter B , the pair reflected from the back of the plate will provide an indication of the changes mentioned above. The intensity of the current due to a single beam is about one-quarter of the peak fringe intensity, and, if the single beams are to be used as a reference, it is desirable that the fringe currents at the working position should also be of this order of magnitude. The change in the amplifier current when the unknown fluid is substituted for the standard can then be corrected by subtracting the corresponding change with one pair of the beams blocked off by the shutter. It is clear that this method cannot be used if the fluids differ appreciably in absorption. The change in amplifier current can be calibrated to read directly small shifts of the fringes, an advantage if the refractive index is varying rapidly.

It should be mentioned that the presence of the plates F' in the lower beam causes the distance between neighbouring fringes in the focal plane of the lens L_2 to be slightly different in the two systems. For the very small angular separation of the plates F' used here this effect is barely noticeable. Both the methods of measurement described above ensure that the difference in fringe spacing has no effect on the readings. For visual observations, however, the plates F' will usually be at a larger angle and the difference in the fringe spacing may be quite marked. This difference in spacing may be eliminated by placing in the upper beam a pair of plates similar to the pair F' in the lower beam.

V. PERFORMANCE OF THE REFRACTOMETER

(a) Calibration of the Jamin Compensator

It is first necessary to determine the displacement of the fringes produced by the fringe-shifting plates F' . The displacement for a Jamin compensator is not strictly proportional to the angle of incidence i , but varies according to the expression (Hansen 1930)

$$\sin i \left[1 - \frac{\cos i}{(\mu^2 - \sin^2 i)^{\frac{1}{2}}} \right], \quad \dots \dots \dots (1)$$

where μ is the refractive index of the glass. Thus as i increases the rotation of the compensator plates required to produce a given fringe-shift decreases. In the present equipment the rotation is produced by a micrometer tangent screw bearing against a pivoted bar, and so, if x is the screw setting and L the length of the bar, the rotation is equal to $\tan^{-1}x/L$. A maximum shift of a little more than one-half of a fringe in each direction is provided, the integral part of the fringe-shift being assumed known from less accurate measurements.

It is preferable to calibrate the compensator rather than to rely on expression (1). The shaft carrying the plates is arranged so that it can be rotated independently of the tangent screw. Thus the rotation necessary to shift through

one fringe (e.g. from one minimum to the next) can be determined for the four fringes surrounding the working position, and, using finite differences, a formula can then be found for the fringe-shift in terms of the rotation. This can then be converted into a formula giving the fractional shift f in terms of the screw movement x . Fortunately, the non-linearity introduced by the expression (1) largely cancels that introduced by the inverse tangent term. The final expression for the fractional fringe-shift f is found in the present apparatus to be

$$f = \frac{x}{45.683}(1 - 0.00179f). \quad \dots\dots\dots (2)$$

Since x can be read to better than 0.003, the sensitivity is better than 1/15,000th of a fringe.

This expression was determined at a temperature of 23 °C, most of the measurements being made in the vicinity of this temperature. If x is large and great accuracy is required, it may be necessary to recalibrate the screw at the working temperature. For measurements on gases the pressure may be adjusted so that x is always small.

The screw may also be calibrated by varying the pressure in a short air cell.

(b) Stability

A stability of better than 1/5000th of a fringe over a period of 20–30 min is claimed for the instrument. Table 1 shows a typical series of readings of the setting x of the tangent screw at which the fringes coincide. These were taken with liquid cells containing water, 1.3 cm in length, in the interfering beams.

TABLE 1
STABILITY OF THE REFRACTOMETER

Time	Tangent Screw Reading, x	Drift Between Readings	
		$\Delta x (\times 10^3)$	$\Delta f (\times 10^4)^*$
1420	0.437		
1440	0.443	+6	+1.3
1500	0.447	+4	+0.9
1520	0.440	—7	—1.5
1540	0.443	+3	+0.7
1600	0.436	—7	—1.5
1620	0.430	—6	—1.3
1640	0.437	+7	+1.5
1700	0.436	—1	—0.2
1720	0.428	—8	—1.8
1740	0.430	+2	+0.4

* Calculated from equation (2), Section V (a).

For the readings in Table 1 the maximum deviation from the mean of the set (0.437) over the period of 200 min corresponds to a change in f of 2.2×10^{-4} . The long-term stability is not always as good as this if the room temperature is varying rapidly.

(c) Reproducibility of Measurements with Liquids

The design of the refractometer is not ideally suited to the measurement of liquid indices. The lid of the box shielding the refractometer from draughts has to be removed and then the cell cover, when the liquid is being changed. The liquids are inserted and removed by a dropper. It is necessary to wait for a period of 10–20 min to allow the system to return to temperature equilibrium. A separate refractometer for liquids with much shorter beam paths would probably give better results.

It appears that the readings of x are reproducible to better than 1/2000th of a fringe. Table 2 shows a set of successive values of x obtained by removing the liquid from the cell and inserting fresh liquid from the same sample.

TABLE 2
DRIFT ON REPLACING LIQUID

Time	Tangent Screw Reading, x	Drift	
		$\Delta x (\times 10^3)$	$\Delta f (\times 10^4)^*$
1750	0.884		
1805	0.880	—4	—0.9
1820	0.883	+3	+0.7
1835	0.887	+4	+0.9
1850	0.904	+17	+3.7
1905	0.912	+8	+1.8

* Calculated from equation (2), Section V (a).

(d) Refractive Index Differences for Liquids

The refractive index difference $\Delta\mu$ is given by $(\lambda/t)\Delta f$, where λ is the wavelength and t the cell thickness. On substituting the values $\lambda = 0.54623 \times 10^{-4}$ cm and $t = 1.3080$ cm, and using equation (2),

$$\Delta\mu = 91.41 \times 10^{-8} \Delta x. \quad (3)$$

It is usually best to take a set of three readings—standard, unknown, standard—the central reading being then subtracted from the mean of the other two. This guards against possible errors due to disturbance of the apparatus or contamination of the liquids, and also compensates for drifts of the fringes which may occur if there are large temperature gradients present. To obtain greater accuracy a more extended set of readings may be taken. Table 3 shows a typical set for two samples of water (labelled A and B).

The successive differences are not independent, but the mean is the most easily calculated estimate and appears to be reasonably efficient. An estimate s_m of the standard error of the mean can be found from the residuals v , the formula being

$$s_m^2 = \left[\frac{4(4n-11)}{6n^3-52n^2+148n-136} \right] \Sigma v^2,$$

where n is the original number of readings. A discussion of the statistics would be out of place here.

A good check on the reliability of the measurements is afforded by the mutual intercomparison of three liquids. For three water samples identified as A, B, and C, the following differences were obtained:

$$A-B, 10^8\Delta\mu=28.7\pm0.3;$$

$$B-C, 10^8\Delta\mu=24.8\pm0.7;$$

$$C-A, 10^8\Delta\mu=3.5\pm0.8.$$

The general experience has been that the error in a single determination of refractive index difference of this magnitude is almost always less than 2×10^{-8} , while a more extended set of observations enables the difference to be determined to at least 1×10^{-8} .

TABLE 3
REFRACTIVE INDEX DIFFERENCE FOR LIQUIDS A AND B

Time	Liquid	Tangent Screw Reading, x	Mean of Adjoining x	$\Delta x (\times 10^3)$	$\Delta\mu (\times 10^8)^*$
1650	A	0.533			
1705	B	0.230	0.545	31.5	28.8
1720	A	0.557	0.240	31.7	29.0
1735	B	0.250	0.563	31.3	28.6
1750	A	0.569	0.261	30.8	28.2
1805	B	0.272	0.590	31.8	29.1
1820	A	0.612			
Mean	31.4	28.7
Standard error	0.3	0.3
Maximum deviation from mean	0.6	0.5

* Calculated from equation (3), Section V (d).

(e) The Refractometer for Gases

The cell length is 20.74 cm, and so for the mercury green line

$$\Delta\mu=263.4\times10^{-8}\Delta f. \quad (4)$$

In terms of the change in the scale reading x ,

$$\Delta\mu=5.765\times10^{-8}\Delta x. \quad (5)$$

(f) Changes in Atmospheric Pressure

The refractometer will give a direct indication of short-time variations in the atmospheric pressure. If the refractive index is μ at the pressure p ,

$$\Delta\mu=\frac{\mu-1}{p}\Delta p.$$

For standard dry air at $t^\circ\text{C}$,

$$10^3\Delta p=(149.4+0.55t)\Delta x,$$

the pressure change being measured in millimetres of mercury. Assuming a limit of 0.01 for Δx , the limit of reading for the pressure change is about 2

microns. If high accuracy is required the humidity of the air should be allowed for. The values obtained for Δp agree with those given by an oil manometer to the limit of reading of the latter (about 10 microns).

The refractometer will also follow very rapid changes in pressure. Its use as a microphone, with a speaking tube attached to one of the cells, provides an interesting novelty. When the EF55 cathode voltage is fed through a power amplifier to a loudspeaker quite good reproduction is obtained, although the noise from the photomultiplier tubes is rather high.

(g) *Measurement of Refractive Indices of Gases*

Since it is possible to set the scale reading x to well within 0.01 , it should be possible to read the refractive index of a gas to within 5×10^{-10} . However, such an accuracy certainly cannot be obtained outside a standards laboratory, since it requires a very accurate knowledge of the pressure and temperature and of the cell length. Even the dispersion equations for air obtained by the principal standards laboratories give values of refractive index differing by up to 1×10^{-7} (Barrell 1951).

TABLE 4
REFRACTIVE INDEX DIFFERENCE FOR AIR AND COMMERCIAL OXYGEN
Mean temperature, 23.0°C ; mean pressure, 757.5 mm Hg

Time	Difference in Screw Reading, Δx	Difference in Refractive Index, $\Delta\mu(\times 10^6)^*$
1345	-3.01	18.61
1425	-0.90	18.49
1450	+0.83	18.39
1520	-0.93	18.49
1545	-1.80	18.54
Mean		18.50
Standard error		0.04
Maximum deviation from mean ..		0.11

$$* 10^6\Delta\mu = 7 \times 2.634 - 0.0576\Delta x.$$

It seems possible without any elaborate equipment to compare the refractive index of a gas with that of standard air to an accuracy of about 5×10^{-8} . Table 4 gives a series of readings of the change Δx when oxygen (of commercial quality as used for glass-blowing etc.) was substituted for dry air in one of the cells. The value of the integral order—in this case 7—was calculated from refractive index tables. This value was later checked by direct counting of the oscillations of the meter needle as the gases were removed.

The values $\Delta\mu$ show a systematic trend, due at least in part to variations in the atmospheric pressure and temperature, and perhaps also to slight changes in the constitution of the air in the room. The mean value of $10^6\Delta\mu$ was 18.50 ± 0.04 . When this is reduced to standard conditions (15°C , 760 mm),

the value 19.08 ± 0.04 is obtained. Using the value $(277.93 \pm 0.03) \times 10^{-6}$ for the refractivity of standard air, the refractivity of the oxygen sample is then $(258.85 \pm 0.05) \times 10^{-6}$.

VI. ACKNOWLEDGMENTS

The author would like to express his thanks to Mr. J. Davies of the Physics Workshop, who constructed the refractometer, and to his laboratory assistant, Mr. S. W. Hutchinson, who constructed the associated equipment.

VII. REFERENCES

- BARRELL, H. (1951).—*J. Opt. Soc. Amer.* **41** : 295-9.
GUEST, P. G., and SIMMONS, W. M. (1953).—*J. Opt. Soc. Amer.* **43** : 319-20.
HANSEN, G. (1930).—*Z. InstrumKde.* **50** : 460-74.
INGELSTAM, E. (1953).—*Ark. Fys.* **6** : 287-316.
SCROGGIE, M. G. (1952).—*Wireless World* **58** : 14-8.

THE EFFECTS OF COLLIMATION AND OBLIQUE INCIDENCE IN LENGTH INTERFEROMETERS. I

By C. F. BRUCE*

[*Manuscript received November 22, 1954*]

Summary

An investigation has been made of the effects of position and size of diaphragm apertures on the interference fringes in length interferometry. The results indicate that, while the well-known correction formula for the effect of the oblique rays from a point source off the optic axis is always applicable, the effect of finite area of aperture is a more complex one and is not in general proportional to the length measured. This effect depends essentially on the phase differences arising from different points of the aperture and is oscillatory in character. For a narrow slit the maximum fringe displacement arising from this effect is not greater than about 0.15 fringe. For a circular aperture the maximum effect is about 0.5 fringe and for a square aperture about 0.35 fringe.

I. INTRODUCTION

In precision length interferometry, measurements are made most frequently with Fizeau-type and Michelson-type interferometers (Kösters 1926; Pérard 1930; Barrell 1948). In all cases collimation of light is obtained by placing the entrance aperture at the principal focus of the collimating lens. Since the aperture has a finite area the emergent beam from the collimating lens is not ideally collimated. In the case of Fizeau systems the aperture is also very frequently placed off the optic axis so that the light as a whole falls obliquely on to the interferometer. The effects of oblique incidence on the disposition of the fringes is well known and their significance is becoming of greater importance as isotopic light sources extend the range of direct measurement by interferometry.

Figures 1 and 2 show the optical schemes of the National Physical Laboratory Gauge Interferometer and the Kösters Gauge Interferometer which are Fizeau and Michelson systems respectively. In the Fizeau system the illuminating (entrance) and viewing (exit) apertures are off the optic axis so that the effects of both oblique incidence and area of aperture are involved. In the Kösters interferometer both entrance and exit apertures lie on the optic axis so that any effect from oblique rays can only come from the resultant effect of all rays from the area of the aperture. For convenience the effect of aperture area alone can be termed a "collimation effect" and the effect of non-normal incidence an "incidence effect". In both cases of course the effects arise from oblique rays.

The correction factors for the "incidence effect" are derived from the classical relation $n\lambda = 2t \cos \theta$ where θ is the angle of incidence and t corresponds to the length measured. The correction factor per unit length is $(1 - \cos \theta)$

* Division of Metrology, C.S.I.R.O., University Grounds, Sydney.

or $\frac{1}{2}\theta^2$ for small values of θ . If in the case of the Fizeau system the distance between entrance and exit apertures is $2s$, and f is the focal length of the collimating lens, the correction factor for the incidence effect is

$$C_1 = s^2/2f^2. \quad \dots\dots\dots (1)$$

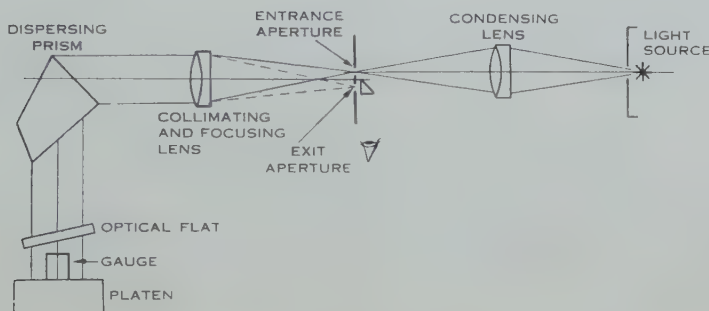


Fig. 1.—Optical scheme of N.P.L. gauge interferometer.

The effect of area of aperture has in the past been considered to be the integrated sum of obliquity effects as determined by equation (1) for all elementary point sources within the area of the aperture. Thus for a circular aperture of radius r (Fig. 3) the correction would be

$$C_2 = \frac{\int_0^{2\pi} \int_0^r \frac{x^2}{2f^2} x d\phi dx}{\int_0^{2\pi} \int_0^r x d\phi dx} = \frac{r^2}{4f^2}.$$

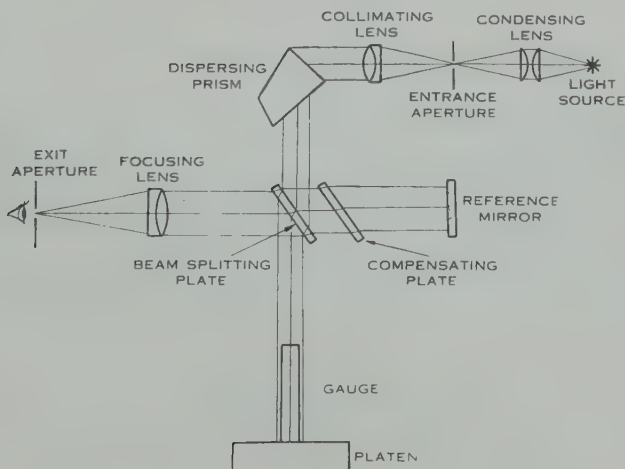


Fig. 2.—Optical scheme of Kösters gauge interferometer.

For a rectangular aperture of length l and breadth b , $C_2 = (l^2 + b^2)/24f^2$ Hamon (1951). The total correction for circular apertures in the Fizeau system would be

$$C = C_1 + C_2 = s^2/2f^2 + r^2/4f^2. \quad \dots\dots\dots (2)$$

For the Kösters interferometer with rectangular apertures, the total correction would be

$$C_2 = (l^2 + b^2)/24f^2. \quad \dots\dots\dots (3)$$

The above integration does not take into account the actual intensity distribution of the fringes and it is considered that the actual observed shape and position of a fringe depends on the integrated intensity from all elementary fringes formed from all elementary sources within the entrance aperture. It should be stated, of course, that the dimensions of the entrance aperture will control the effect from its area provided its image, focused in the plane of the exit aperture, is smaller than the size of the exit aperture. If the size of the image of the entrance aperture is greater than the size of the exit aperture, then clearly the dimensions of the latter and the focal length of the focusing lens associated with it will determine the range of oblique rays collected by the observer's eye.

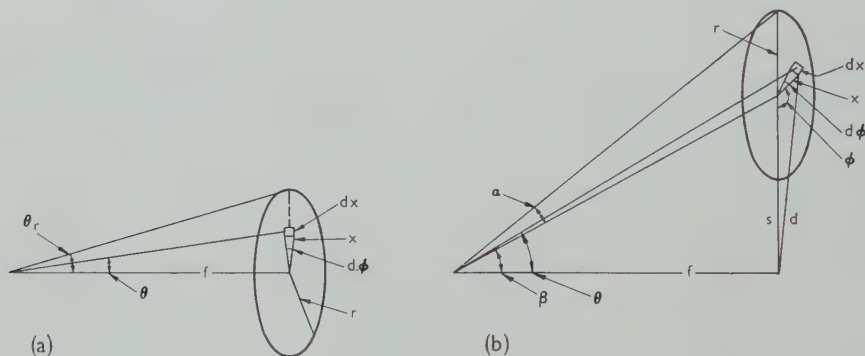


Fig. 3.—Theory of fringe intensity. (a) Circular aperture on centre, (b) circular aperture off centre.

It is the purpose of this paper (paper I) to describe an investigation of the effects of oblique incidence and area of aperture and the accuracy of the above relationships. Paper II (Thornton 1955), which follows this paper, describes the theory of the effects in greater detail for the various possible cases in both two-beam and multiple-beam systems. It is assumed that fringe displacements are observed from the centre of fringes since there was no real evidence that settings are made on different parts of different sets of fringes as a result of fringe asymmetry.

II. FRINGE INTENSITY DISTRIBUTIONS

The intensity for two-beam fringes is given by $I = \cos^2 K$ where $K = \frac{1}{2}$ (phase difference) and where for simplicity the amplitude factor is treated as unity.

(a) Circular Aperture

For a circular aperture centred on the optic axis (Fig. 3 (a)) the intensity for an elementary point source making an angle θ with the axis is

$$\delta I = \cos^2 (K \cos \theta) x d\phi dx.$$

The resultant fringe intensity due to the whole aperture is

$$I = \int_0^{2\pi} \int_0^r \cos^2 (K \cos \theta) x d\phi dx.$$

As θ is of the order of 0.001 radian a legitimate approximation for $\cos \theta$ is $1 - x^2/2f^2$ where f is the focal length of the collimating lens as indicated in Figure 3. The integral then becomes

$$I = 2\pi \int_0^r \frac{1}{2} [\cos 2K(1 - x^2/2f^2) + 1] x dx.$$

By substituting $u = 2K(1 - x^2/2f^2)$ and integrating

$$I = \frac{\pi r^2}{2} - \frac{\pi f^2}{2K} (\sin 2K \cos Kr^2/f^2 - \cos 2K \sin Kr^2/f^2 - \sin 2K). \quad \dots (4)$$

A convenient form of equation (4) is

$$I = \frac{1}{2} A \left[1 + \frac{\sin \Delta}{\Delta} \cos (2K - \Delta) \right], \quad \dots \dots \dots (5)$$

where $A = \pi r^2$, which is the area of the aperture, and $\Delta = \frac{1}{2} K \theta^2$, $\theta = r/f$, $K = 2\pi t/\lambda$, and t is the length measured. The phase factor Δ is the phase difference between a point source at the centre and a point source at the edge of the aperture. The relation indicates that the fringe would be symmetrical and the phase factor Δ determines where fringe minima (or maxima) will occur. If Δ is small $\sin \Delta/\Delta$ is approximately unity and equation (5) becomes

$$I = A \cos^2 (K - \frac{1}{2}\Delta), \quad \dots \dots \dots (6)$$

which shows that fringe minima will be displaced from the position of those for a point source on the axis by $\frac{1}{2}\Delta$. It can be seen, however, from equation (5) that for $0 \leq \Delta \leq \pi$, the factor $\sin \Delta/\Delta$ is positive throughout and the fringe displacement will still be given by $\frac{1}{2}\Delta$.

For values of Δ greater than π , the effects are different due to the change in sign of $\sin \Delta/\Delta$. At $\Delta = n\pi$ ($n=1, 2, 3 \dots$) the intensity is the same for all values of the phase factor $\cos (2K - \Delta)$ and the fringes should therefore vanish for these values of Δ . For the range $\pi \leq \Delta \leq 2\pi$ the intensity is

$$\begin{aligned} I &= \frac{1}{2} A \left[1 + \frac{\sin (\pi + \delta)}{\pi + \delta} \cos (2K - \pi - \delta) \right] \\ &= \frac{1}{2} A \left[1 + \frac{\sin \delta}{\pi + \delta} \cos (2K - \delta) \right], \end{aligned}$$

where $\Delta = \pi + \delta$ and $0 \leq \delta \leq \pi$. The fringe displacement is therefore $\frac{1}{2}\delta$ or $\frac{1}{2}(\Delta - \pi)$.

In general for the range $n\pi \leq \Delta \leq (n+1)\pi$ the fringe displacement will be $\frac{1}{2}\delta$ or $\frac{1}{2}(\Delta - n\pi)$ where $\Delta = n\pi + \delta$ and the maximum fringe displacement will not

exceed $\frac{1}{2}\pi$ or 0.5 fringe. The form of the obliquity-fringe displacement curve will therefore be as shown in curve A of Figure 4.

As the fringes vanish for $\Delta = n\pi$ their visibility is of some interest. The "visibility" is usually defined as

$$V = \frac{I_{\max.} - I_{\min.}}{I_{\max.} + I_{\min.}}$$

The maximum and minimum values of intensity will be at $2K - \Delta = 2n\pi$ and $(2n+1)\pi$ respectively, for $0 \leq \Delta \leq \pi$, at $2K - \Delta = (2n+1)\pi$ and $2n\pi$ respectively, for $\pi \leq \Delta \leq 2\pi$, and so on. The visibility V will be given by $\sin \Delta / \Delta$ and the maximum value for V is given by $dV/d\Delta = 0$ which gives the condition $\Delta = \tan \Delta$

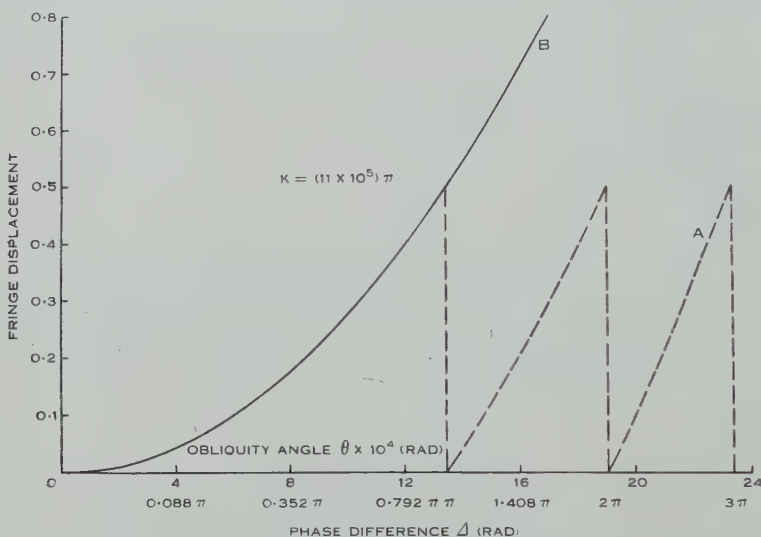


Fig. 4.—Fringe displacement-obliquity curve for a circular aperture. Curve A from $I = \frac{1}{2}\Delta[1 + (\sin \Delta / \Delta) \cos (2K - \Delta)]$, curve B from $C = d^2/16f^2$ ($d = 2r$).

for best visibility. Thus for very small values of Δ the maximum value for visibility will approach 1. For $\pi \leq \Delta \leq 2\pi$ the visibility maximum will be about 0.2 with Δ near $3\pi/2$ and for $2\pi \leq \Delta \leq 3\pi$ it will be about 0.1 with Δ near $5\pi/2$. Fringes are very faint but visible for values of Δ up to 5π at least in the experimental work.

For a circular aperture off centre (Fig. 3 (b)) the intensity will be

$$\begin{aligned} I &= \int_0^{2\pi} \int_0^r \cos^2 (K \cos \theta) x dx d\varphi \\ &= \int_0^{2\pi} \int_0^r \frac{1}{2} [1 + \cos (2K - Kd^2/f^2)] x dx d\varphi, \end{aligned}$$

where $\theta = d/f$ and $d^2 = x^2 + s^2 + 2xs \cos \varphi$.

Substituting and expanding,

$$I = \frac{1}{2}\pi r^2 + \pi \cos 2K \int_0^r J_0(z) \cos \frac{K(x^2+s^2)}{f^2} x dx \\ + \pi \sin 2K \int_0^r J_0(z) \sin \frac{K(x^2+s^2)}{f^2} x dx,$$

where $J_0(z)$ is the zero order Bessel function in which

$$z = 2Kxs/f^2.$$

The intensity can therefore be represented by

$$I = \frac{1}{2}\pi r^2 + \pi \int_0^r \cos \left[2K - \frac{K(x^2+s^2)}{f^2} \right] J_0(2Kxs/f^2) x dx. \dots (7)$$

As xs/f^2 is of the order of 10^{-6} , the function $J_0(2Kxs/f^2)$ would approach unity only if K is less than, for example, 10^5 . With this limitation equation (7) becomes

$$I = \frac{1}{2}\pi r^2 + \pi \int_0^r \cos \left[2K - \frac{K(x^2+s^2)}{f^2} \right] x dx,$$

which reduces to

$$I = \frac{\pi r^2}{2} + \frac{\pi f^2}{K} [\sin \frac{1}{2}K\alpha^2 \cos (2K - \frac{1}{2}K\alpha^2 - K\beta^2)],$$

where $\alpha = r/f$ and $\beta = s/f$.

This equation can be put in the following form

$$I = \frac{1}{2}A \left[1 + \frac{\sin \Delta}{\Delta} \cos (2K - \Delta - K\beta^2) \right], \dots (8)$$

where $\Delta = \frac{1}{2}K\alpha^2$.

This equation is similar to equation (5) and for $\Delta \gg \pi$ reduces to

$$I = A \cos^2 (K - \frac{1}{2}\Delta - \frac{1}{2}K\beta^2). \dots (9)$$

Thus the fringe displacement would be $\frac{1}{2}\Delta + \frac{1}{2}K\beta^2$, which is equivalent to a normal obliquity effect from the centre of the aperture ($\frac{1}{2}K\beta^2$) plus the effect of area of aperture as found for the centred case. This result is not exact in view of the approximation $J_0(z) = 1$ which is not legitimate for higher values of K , but it is some indication of the conditions for smaller values of K .

(b) *Narrow Slit*

The intensity distribution can be examined by a summation of intensities from elementary point sources making up the aperture and can be done graphically or analytically. This method can be used conveniently for the case of a narrow slit which is generally used in the Kösters interferometer.

The intensities for a number of point sources $P_0, P_1, P_2 \dots$ (Fig. 5) can be expressed as

$$\begin{aligned} I_0 &= \cos 2K, \\ I_1 &= \cos^2 (K \cos \theta), \\ I_2 &= \cos^2 (K \cos 2\theta), \\ I_n &= \cos^2 (K \cos n\theta). \end{aligned}$$

The condition for fringe minima can be found by differentiating these expressions with respect to K and equating the result to zero. As $n\theta$ is always small, terms involving θ^2 as factor may be neglected and the result becomes

$$\begin{aligned} \sin 2K(1 + \cos 2\varphi + \cos 8\varphi + \dots + \cos 2n^2\varphi) \\ = \cos 2K (\sin 2\varphi + \sin 8\varphi + \dots + \sin 2n^2\varphi), \end{aligned}$$

where $\varphi = \frac{1}{2}K\theta^2$.



Fig. 5.—Theory of fringe intensity for a rectangular aperture.

The values of K at which there are fringe minima or maxima is therefore given by

$$\tan 2K = \frac{\sum_1^n \sin 2n^2\varphi}{1 + \sum_1^n \cos 2n^2\varphi} + n\pi. \quad \dots\dots\dots (10)$$

Values of K derived from expressions of this type for a range of values of n agree closely with those obtained from direct graphical summation of intensities. The general character of the obliquity-fringe displacement curves are shown in Figures 6 (a) and (b) for values of K of about $(4 \times 10^5)\pi$ and $(11 \times 10^5)\pi$ respectively. For $\lambda = 5461 \text{ \AA}$ these values would correspond to lengths of about 110 mm and 300 mm respectively. Curve B is obtained by using equation (3) with $b=0$.

The curves indicate that the maximum fringe displacement possible is about 0.15 fringe and up to this value the displacements are approximately of the same order as that derivable from equation (3) if b is taken to be zero in comparison with l . This maximum occurs when Δ is about $\frac{1}{2}\pi$ and thereafter there is no agreement with equation (3) as the fringe displacement is oscillatory in character. It is of interest to note that the phase value of $\Delta = \frac{1}{2}\pi$ corresponds to the approximate limiting condition (Rayleigh criterion) for no deterioration in the visibility of the fringes. The fringes of course are clearly observable at greater values of this phase difference. For long lengths, phase differences greater than $\frac{1}{2}\pi$ may very well occur if the focal length of the lens is not sufficiently

great or the dimensions of the aperture are not sufficiently small. For example, with a 500 mm length and $\lambda = 5000 \text{ \AA}$, K is about $(2 \times 10^6)\pi$, and if the maximum obliquity angle θ was no greater than 0.001 radian the phase factor $\Delta = \frac{1}{2}K\theta^2$ would be π radians.

An exact treatment of the centred narrow slit case is given in paper II in terms of Fresnel integrals and shows that the curves given here indicate adequately the oscillatory character of the fringe displacement and its order of magnitude. Using Fresnel integrals, the conditions for a narrow slit off the axis can be examined to see if the fringe displacement effect can be treated as a normal obliquity effect from the centre of the slit plus the effect of the area of the slit treated as if it were centred on the axis.

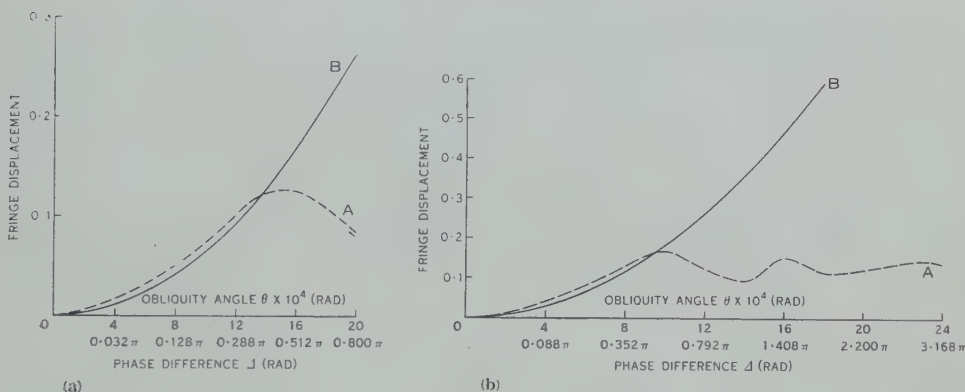


Fig. 6.—Fringe displacement-obliquity curves for a narrow slit. Curve A from summation of elements, curve B from $C = l^2/24f^2$. (a) $K = (4 \times 10^5)\pi$; (b) $K = (11 \times 10^5)\pi$.

For the narrow slit off centre the relevant part of the intensity relation is

$$I = \int_{\theta_1}^{\theta_2} \cos^2 (K \cos \theta) d\theta = \int_{\theta_1}^{\theta_2} \cos^2 K(1 - \frac{1}{2}\theta^2) d\theta,$$

where θ_1 and θ_2 are the obliquity angles at the extreme edges of the slit relative to the optic axis. Differentiating under the integral sign, as done in paper II, with respect to K and ignoring terms with θ^2 as factor, since θ is small

$$\begin{aligned} \frac{\partial I}{\partial K} &= \int_{\theta_1}^{\theta_2} \sin (2K - K\theta^2) d\theta \\ &= \int_0^{\theta_2} \sin (2K - K\theta^2) d\theta - \int_0^{\theta_1} \sin (2K - K\theta^2) d\theta \\ &= \sin 2K \left[\int_0^{\theta_2} \cos K\theta^2 d\theta - \int_0^{\theta_1} \cos K\theta^2 d\theta \right] \\ &\quad - \cos 2K \left[\int_0^{\theta_2} \sin K\theta^2 d\theta - \int_0^{\theta_1} \sin K\theta^2 d\theta \right]. \end{aligned}$$

The condition for fringe minima is therefore

$$\tan 2K = \frac{S(v_2) - S(v_1)}{C(v_2) - C(v_1)}, \quad \dots\dots\dots (11)$$

where $S(v)$ and $C(v)$ are the Fresnel integrals

$$\int_0^v \sin \frac{1}{2}\pi v^2 dv \text{ and } \int_0^v \cos \frac{1}{2}\pi v^2 dv$$

respectively, in which $v = \sqrt{(2K/\pi)\theta}$. The fringe displacement for typical values of θ_1 and θ_2 can be found by evaluating the integrals for the appropriate values of v_1 and v_2 .

It was also shown in the case of the centred narrow slit that there was reasonable agreement between the fringe displacement calculated from equation (3) with $b=0$ and that obtained from the summation of elements analysis described, provided the phase difference between the centre and edge of the slit did not exceed about $\frac{1}{2}\pi$. Accepting this result, the fringe displacement from the effect of area of aperture in the off-centre case will be $\frac{1}{6}K(\theta_2 - \theta)^2$ provided $\frac{1}{2}K(\theta_2^2 - \theta^2)$ is not greater than about $\frac{1}{2}\pi$ where $\theta = \theta_1 + \frac{1}{2}\theta_2$ is the obliquity angle corresponding to the centre of the slit. The fringe displacement due to obliquity of the centre of the slit will be $\frac{1}{2}K\theta^2$. The combined effect is therefore

$$D = \frac{1}{2}K\theta^2 + \frac{1}{6}K(\theta_2 - \theta)^2. \quad \dots\dots\dots (12)$$

Fringe displacements calculated from equations (11) and (12) for typical values of θ_1 , θ_2 , and K agree to well within 0.05 fringe. This indicates again that an off-centre aperture can be treated without great error as a centred aperture for the area effect together with a direct obliquity effect from the centre of the aperture. Both effects are, of course, included in equation (11) which can be used to obtain directly the fringe displacement arising from both oblique incidence and collimation effects.

(c) Rectangular Aperture

For a rectangular aperture centred on the axis whose width may be comparable with its length the intensity is given by

$$\begin{aligned} I &= \int_{-l/2}^{l/2} \int_{-b/2}^{b/2} \cos^2 (K \cos \theta) dy dx \\ &= \int_{-l/2}^{l/2} \int_{-b/2}^{b/2} \frac{1}{2}(1 + \cos 2K \cos \theta) dy dx. \end{aligned}$$

Differentiating the relevant portion with respect to K and ignoring terms with θ^2 as factor

$$\begin{aligned} \frac{\partial I}{\partial K} &= \int_{-l/2}^{l/2} \int_{-b/2}^{b/2} \sin (2K - K\theta^2) dy dx. \\ &= 2 \sin 2K \int_0^{l/2} \int_0^{b/2} \cos K\theta^2 dy dx - 2 \cos 2K \int_0^{l/2} \int_0^{b/2} \sin K\theta^2 dy dx. \end{aligned}$$

Fringe minima (or maxima) will occur when $\partial I/\partial K=0$ which gives

$$\tan 2K = \frac{\int_0^{l/2} \int_0^{b/2} \sin K\theta^2 dy dx}{\int_0^{l/2} \int_0^{b/2} \cos K\theta^2 dy dx}.$$

Making the substitution $\theta^2 = x^2 + y^2/f^2$ where f is the focal length of the lens concerned

$$\begin{aligned} \int_0^{l/2} \int_0^{b/2} \sin K\theta^2 dy dx &= \int_0^{l/2} \int_0^{b/2} \sin \frac{Kx^2}{f^2} \cos \frac{Ky^2}{f^2} dy dx \\ &+ \int_0^{l/2} \int_0^{b/2} \cos \frac{Kx^2}{f^2} \sin \frac{Ky^2}{f^2} dy dx \\ &= f^2 \frac{\pi}{2K} [S(u)C(v) + C(u)S(v)], \end{aligned}$$

where $S(u)$, $C(u)$, $S(v)$, and $C(v)$ are the Fresnel integrals

$$\int_0^u \sin \frac{1}{2}\pi u^2 du, \quad \int_0^u \cos \frac{1}{2}\pi u^2 du, \quad \int_0^v \sin \frac{1}{2}\pi v^2 dv, \quad \text{and} \quad \int_0^v \cos \frac{1}{2}\pi v^2 dv$$

respectively, in which

$$\begin{aligned} u &= \sqrt{\left(\frac{2K}{\pi}\right)} \frac{x}{f}, \\ v &= \sqrt{\left(\frac{2K}{\pi}\right)} \frac{y}{f}. \end{aligned}$$

In a similar manner it can be shown that

$$\int_0^{l/2} \int_0^{b/2} \cos K\theta^2 dy dx = f^2 \frac{\pi}{2K} [C(u)C(v) - S(u)S(v)].$$

The condition for fringe minima is therefore given by

$$\tan 2K = \frac{S(u)C(v) + C(u)S(v)}{C(u)C(v) - S(u)S(v)}. \quad \dots\dots\dots (13)$$

If the width of the aperture is very small compared with its length v will be very small and $S(v) \rightarrow 0$ and the equation becomes

$$\tan 2K = \frac{S(u)}{C(u)},$$

which is the case of the narrow slit as derived in paper II.

The fringe displacements can be found by direct evaluation of the Fresnel integral products in equation (13) for different length/width ratios. As an example Figure 7 shows the fringe displacement curve for a square aperture ($u/v=1$) and for a narrow slit. Values of $u = \sqrt{\left(\frac{2K}{\pi}\right)}\theta$ and $\Delta = \frac{1}{2}K\theta^2$ are shown

as abscissae. The obliquity angle θ refers to that at the edge of the aperture and Δ is the phase difference from the centre to the edge. For the square aperture $\theta = l/2f - b/2f$, and for the equivalent circular aperture $\theta = \sqrt{(l^2 + b^2)}/2f$.

Displacement curves for values of u/v greater than 1 can be found in a similar manner. Examination of these cases indicates that the fringe displacement curves are of the same general character as those for the narrow slit and rectangular aperture and give values for the fringe displacement that are intermediate between these two cases but which gradually increase in the oscillatory region. For a length/width ratio equal to or greater than 4 ($u/v \geq 4$) the aperture can be treated as a narrow slit and this treatment gives correct fringe displacements to within 0.05 fringe for values of Δ of up to at least 2π .

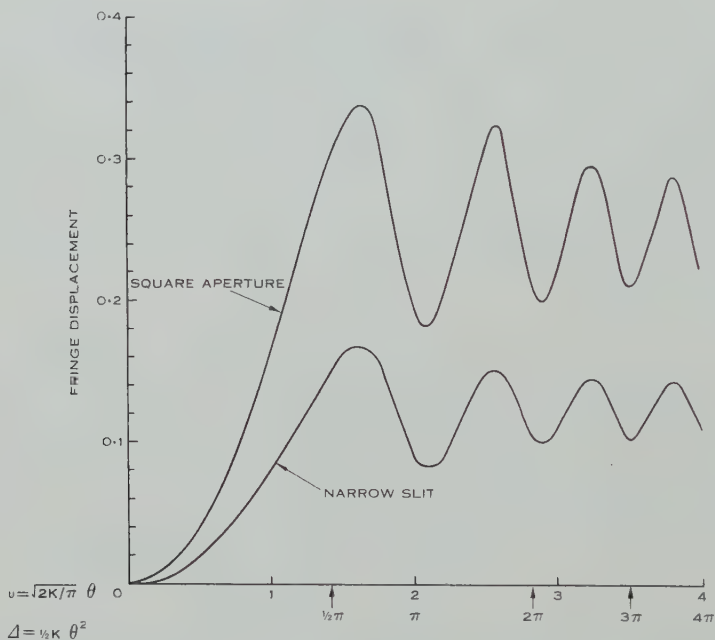


Fig. 7.—Fringe displacement curves for square aperture and narrow slit.

(d) Multiple-Beam Effects

In Fizeau systems it is possible to make use of multiply reflected beams in some degree by coating the optical flat with a highly reflecting layer. Since large air gaps are a necessary condition in practical length measurement full use cannot be made of multiple-beam interference.

The intensity relation for multiple-beam fringes in the centred circular aperture case would be

$$\begin{aligned}
 I &= \int_0^{2\pi} \int_0^r \frac{4R \sin^2 (K \cos \theta)}{(1-R)^2 + 4R \sin^2 (K \cos \theta)} x dx d\varphi \\
 &= 2\pi f^2 \int_0^{\theta_r} \frac{A \sin^2 K(1 - \frac{1}{2}\theta^2)}{1 + A \sin^2 K(1 - \frac{1}{2}\theta^2)} \theta d\theta,
 \end{aligned}$$

where $A=4R/(1-R)^2$. Substituting $u=1-\frac{1}{2}\theta^2$,

$$I=2\pi f^2 \int_{u_0}^{u_r} \frac{A \sin^2 Ku}{-(A+1)+A \cos^2 Ku} du.$$

Integrating,

$$I=2\pi f^2 \left[\frac{1}{K\sqrt{F}} \{ \tan^{-1}(\sqrt{F} \tan K(1-\frac{1}{2}\theta^2)) - \tan^{-1}(\sqrt{F} \tan K) \} + \frac{1}{2}\theta^2 \right].$$

..... (14)

The fringes should therefore be symmetrical, if this was the only effect. The actual effect, however, of a large aperture is to produce a marked asymmetry in the fringes in the direction of increasing order or increasing air gap, for low order, high definition multiple-beam fringes. This is due to the changed phase conditions introduced by the linear displacement of multiply reflected beams at higher angles of incidence (Brossell 1947).

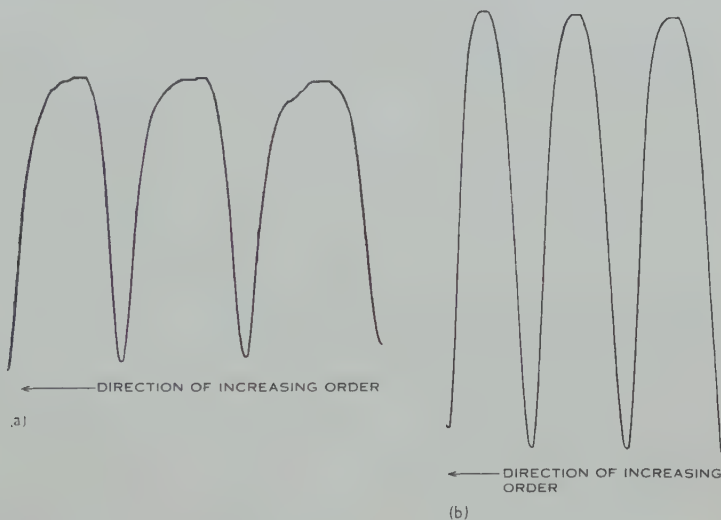


Fig. 8.—Microphotometer record of fringe intensities. (a) Fizeau interferometer, (b) Kösters interferometer.

Microphotometer records were made of the fringe intensity distribution with the interferometers described here at various aperture sizes and at a range of path differences. The aperture size had no observable influence on the fringe shape of the two-beam fringes of the Kösters interferometer or the fringes of the Fizeau interferometer which may be two-beam or partially multiple-beam in character. Typical results are shown in Figures 8 (a) and (b).

It is of interest to note that the Kösters interferometer gave symmetrical fringes and the Fizeau interferometer fringes of smaller half width, but with a distinct asymmetry in the direction of decreasing order, which is the opposite direction to that to be expected from pure collimation effects. This indicates

that phase effects at the silvered optical flat are the predominating influence on the fringe shape and that any effects due to area of aperture are negligible in comparison. The effects of using fringes of the asymmetrical multiple-beam type from the surface of a length standard together with broad two-beam fringes from the platen to which the standard is wrung has been investigated (Terrien 1954).

A treatment of the multiple-beam case by summation of elements for values of the reflection coefficient R in the range from 0.6 to 0.9 gives obliquity-fringe displacement curves of the same form as those for two-beam cases. It should be noted, however, that under true multiple-beam conditions the value of the length t must be very small which in turn makes K much smaller. The fringe displacement factor which is a function of K and θ^2 will therefore be extremely small, and of little practical significance.

III. EXPERIMENTAL OBSERVATIONS

The Kösters interferometer made in this Laboratory permitted the apertures to be varied in size over a small range during observations. For example a particular setting allowed the entrance aperture to be varied from 0.3×0.08 mm to 0.6×0.3 mm. Observations were made initially with an exit (or viewing) aperture 1.0×0.5 mm. As the focal lengths of the lenses associated with the entrance and exit apertures were 208 mm and 390 mm respectively, the image of the entrance aperture was smaller than the size of the exit aperture, with which it is coincident. The dimensions of the entrance aperture therefore controlled the limits on obliquity effects. For a 300 mm length and a focal length of 208 mm a change of 0.4 fringe should occur in the fringe displacement as the entrance aperture changes from 0.3×0.08 mm to 0.6×0.3 mm, if equation (3) is valid. The displacement is that between the fringes from the length gauge surface and the fringes from the surface to which this gauge is wrung. The observed displacement change was certainly no greater than 0.1 fringe.

Further tests were made with a set of circular and rectangular apertures. These were made and used so that their centres were precisely located on the optic axis in order that no off-centre effects would occur. This was further verified by autocollimation tests. A 305 mm (12 in.) length was used and the wavelength 5461 Å from a mercury 198 lamp gave clear fringes at this path difference which corresponds to a value of about $(11 \times 10^5) \pi$ for the phase factor K . The set of apertures was used first at the entrance end of the interferometer and then at the exit end. In each case the fixed aperture, whether it was at the exit end or entrance end, was sufficiently large to ensure that the set of apertures of variable sizes controlled the range of oblique rays received by the observer's eye.

Tables 1 and 2 give a series of results using the exit aperture of the interferometer to control the obliquity due to area of aperture. As the focal length of the associated lens was 390 mm the aperture did not need to be as small as when the entrance aperture was used ($f=208$ mm) to give a range of values for Δ below $\frac{1}{2}\pi$.

It is to be noted that for rectangular apertures there is approximate agreement between the observed fringe displacement difference and those calculated from equation (3) for values up to about $\frac{1}{2}\pi$ only. Thereafter there is disagree-

TABLE 1
RECTANGULAR APERTURES

No.	Dimensions		Obliquity Angle $\theta = \sqrt{(l^2 + b^2)}/2f$ (rad)	Phase Difference $\Delta = \frac{1}{2}K\theta^2$ (rad)	Fringe Displacement Difference	
	l (mm)	b (mm)			From $l^2 + b^2/24f^2$	Observed Mean
1	0.310	0.108	4.20×10^{-4}	0.10π	0	0
2	0.409	0.221	5.96	0.20	0.04	0.02
3	0.644	0.105	8.36	0.39	0.10	0.08
4	0.622	0.332	9.04	0.46	0.12	0.11
5	0.818	0.411	11.73	0.77	0.23	0.15
6	1.030	0.099	13.27	0.98	0.30	0.17
7	1.026	0.611	15.31	1.31	0.41	0.19
8	2.025	0.120	26.01	3.78	1.23	0.13

TABLE 2
CIRCULAR APERTURES

No.	Diameter d (mm)	Obliquity Angle $\theta = d/2f$ (rad)	Phase Difference $\Delta = \frac{1}{2}K\theta^2$ (rad)	Fringe Displacement Difference	
				From $d^2/16f^2$	Observed Mean
1	0.198	2.54×10^{-4}	0.03π	0	0
2	0.408	5.23	0.15	0.06	0.04
3	0.612	7.85	0.34	0.15	0.13
4	0.784	10.05	0.56	0.26	0.25
5	0.990	12.69	0.90	0.43	0.30
6	1.180	15.13	1.28	0.62	0.45
7	1.415	18.14	1.84	0.90	0.53
8	1.512	19.38	2.10	1.02	0.48
9	2.010	25.77	3.71	1.81	0.48
10	2.504	32.10	5.75	2.86	0.48

ment and the observed effects conform quite closely with that indicated by the oscillatory curve of Figures 5 and 6 for a narrow slit.

For the circular apertures the observed displacement generally agrees with equation (2) with $S=0$ for values of Δ up to π . The theoretical analysis

given in this paper also predicts this effect exactly, but for values of Δ greater than π an oscillatory condition again occurs. As Δ approaches π the visibility of the fringes decreases markedly and they are difficult to observe which may

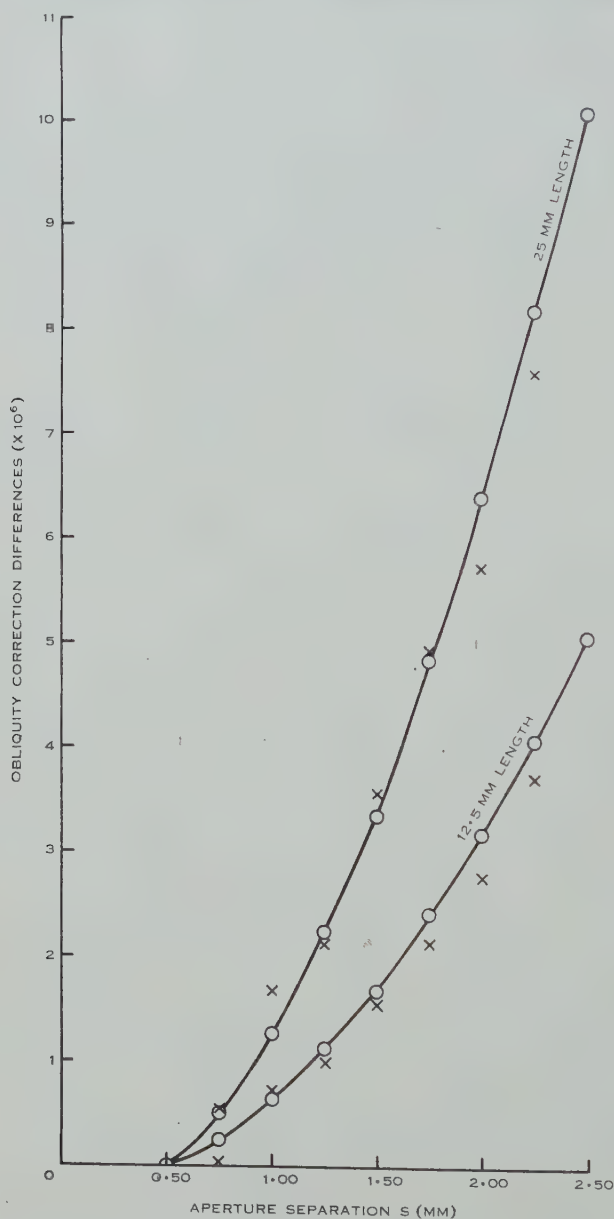


Fig. 9.—Obliquity correction differences for different aperture separations.

account for the rather large disagreement between observed and predicted results for aperture No. 5. For values of Δ greater than π visibility is too low to determine with confidence if the fringe displacement agrees closely with the

predicted oscillatory curve in Figure 4, but there was no evidence of an increase beyond 0.5 fringe.

Experimental tests with the Fizeau interferometer are more difficult since large path differences cannot be used. In addition only a limited separation or enlargement of the apertures can be tolerated if the light beams are to pass through the optical system on their return from the interferometer unit. With increasing aperture separation, the effects of varying obliquity, due to size of aperture as the eye is moved, become more serious.

A series of observations with different apertures were made with lengths from 12.5 to 25 mm corresponding to path differences of 25 and 50 mm. The fringes were sharpened by using a highly reflecting layer on the optical flat. The fringe displacement changed as the aperture separation increased in a manner closely agreeing with equation (1). Figure 9 shows for two lengths obliquity correction differences corresponding to changes in separation of apertures. The full line curve has been drawn from obliquity corrections calculated from equation (1). The observed obliquity correction differences from that for an aperture separation of 0.50 mm are marked \times .

An attempt was made to test the effect of aperture size, but the predicted effects were too small to allow really reliable observation, because of the small path differences and to the limitations mentioned above in the interferometer used. It did not seem profitable to continue such tests when fairly definitive results are possible for centred apertures with the Kösters interferometer.

IV. CONCLUSIONS

The study of the intensity distribution of the fringes together with microphotometer records indicate that the effects of finite area of aperture on symmetry are insignificant. Any asymmetry which arises in the Kösters system is not appreciable and not observable in the case of partially sharpened Fizeau fringes because of the predominating influence of phase effects at the silvered glass surface.

The effect of finite area of aperture on the position of the fringe maximum is a more complex one. For a centred aperture a fringe displacement does occur which progressively increases provided the phase difference Δ between the centre of the aperture and its edge does not exceed a certain value. For narrow slits and rectangular apertures this value is about $\frac{1}{2}\pi$ and for circular holes about π . Beyond these values the fringe displacement does not increase as equations (2) and (3) would indicate but oscillates. While there happens to be approximate agreement theoretically and experimentally with these equations for the limited values of Δ mentioned, there is no evidence that they can be applied generally as corrections per unit length. For example it has been quite common to use a narrow slit in the Kösters interferometer. The analysis described here shows that the maximum fringe displacement possible, due to the size of the slit, will be about 0.15 fringe. Much larger values for the displacement would be possible if a formula of the type exemplified by equation (3) was applied without consideration to the value of the phase factor Δ involved.

A rectangular aperture may be treated as a narrow slit without great error if its length is equal to or greater than four times its width.

In the case of off-centred apertures, fringe displacement effects appear to be adequately covered by treating the effect as one of direct obliquity from the centre of the aperture plus the effect from the area of the aperture treated as a centred aperture. In the case of a narrow slit off centre, the fringe displacement effect owing to both the position and size of aperture can be calculated from an exact relation.

V. ACKNOWLEDGMENTS

The author would like to thank Dr. H. Barrell, of the National Physical Laboratory, England, and M. J. Terrien, of the Bureau International des Poids et Mesures, France, for helpful criticisms. Valuable suggestions and criticisms have also been made by Associate Professor E. O. Hercus of the University of Melbourne. Discussions have also been had with Mr. P. M. Gilet, formerly of this Laboratory, and Mr. M. J. Puttock and Mr. B. S. Thornton of this Laboratory. Assistance with the experimental work and numerical evaluation from Mr. W. A. F. Cuninghame, Mr. J. R. Cleary, Miss I. Herberg, and Miss W. A. Colahan is gratefully acknowledged.

VI. REFERENCES

- BARRELL, H. (1948).—*Research* **1**: 532.
BROSSELL, J. (1947).—*Proc. Phys. Soc. Lond.* **59**: 224.
HAMON, J. (1951).—*Rev. Opt. (Theor. Instrum.)* **30**: 306.
KÖSTERS, W. (1926).—*Z. Feinmech.* **34**: 55.
PÉRARD, A. (1930).—*Trav. Bur. Int. Poids Mes.* **18**: 85.
TERRIEN, J. (1954).—*J. Phys. Radium* **15**: 6.
THORNTON, B. S. (1955).—*Aust. J. Phys.* **8**: 241.

THE EFFECTS OF COLLIMATION AND OBLIQUE INCIDENCE IN LENGTH INTERFEROMETERS. II

By B. S. THORNTON*

[Manuscript received November 22, 1954]

Summary

It is important in length determinations by interferometry to know the effect on the fringe pattern of an entrance aperture of finite size through which the light enters the interferometer. The correction generally applied has been found to be in error (Bruce 1955). This paper is a further discussion of this correction. Both two-beam and multiple-beam systems are discussed and show, in accordance with experiment, that the correction at present applied for the finite size of the aperture appears to be invalid. Formulae are developed and curves are given from which the correction to be applied for any particular length and aperture size can be very easily determined.

I. INTRODUCTION

The usual obliquity correction formula involves a term due to the displacement of the aperture off the optic axis and another due to the finite size of the aperture itself. If the centre of the aperture is displaced a distance x from the optical axis and f is the focal length of the collimator then the angle subtended at the collimator between the aperture and the optic axis is

$$\theta = \tan^{-1} x/f.$$

Now

$$\begin{aligned} p\lambda &= 2t \cos \theta \\ &\simeq 2t(1 - x^2/2f^2). \end{aligned}$$

In this equation p is the fringe order and t is the path difference.

The obliquity is seen to be $x^2/2f^2$ per unit length. There remains the correction for the finite aperture size which is derived by taking the average correction for the whole area of the aperture calculated in the manner above. This correction is therefore given by

$$\frac{\iint \left(\frac{x^2}{2f^2} + \frac{y^2}{2f^2} \right) dx dy}{\iint dx dy}$$

per unit length, and for a circular aperture of diameter d is $d^2/16f^2$ whilst for a rectangular aperture of length a and width b is $(a^2 + b^2)/24f^2$. Since $x=0$ in the Kösters interferometer the only correction applied is the latter. However, in tests made at large path differences with a variable entrance slit, the fringe

* Division of Metrology, C.S.I.R.O., University Grounds, Sydney.

displacements predicted by the usual formula were not observed and further investigation of the effect seemed desirable. The method of investigation presented in this paper involves a different approach to that used in paper I (Bruce 1955) in that differentiation under the integral sign is used and the results expressed in terms of Fresnel integrals.

II. THE FRINGE DISPLACEMENT FOR A FINITE APERTURE

(a) Two-Beam Systems

The intensity distribution of two-beam interference fringes when the finite size of the entrance aperture is considered is easily shown to be

$$I_1 = \int_{-a/2}^{a/2} \int_{-b/2}^{b/2} \cos^2(K \cos \theta) dx dy \simeq bf \int_0^{\theta_r} 2 \cos^2(K \cos \theta) d\theta, \quad \dots (1)$$

for a rectangular aperture of dimensions b in the x direction and a in the y direction ($a \gg b$), $\theta_r = a/2f$, and

$$I_2 = \int_0^{2\pi} \int_0^r \cos^2(K \cos \theta) x dx d\phi = 2\pi f^2 \int_0^{\theta_r} \cos^2(K \cos \theta) \theta d\theta, \quad \dots (2)$$

for a circular aperture of radius r . Also $K = 2\pi t/\lambda$, t is the path difference and x , θ and ϕ , and r are as in Figure 1.

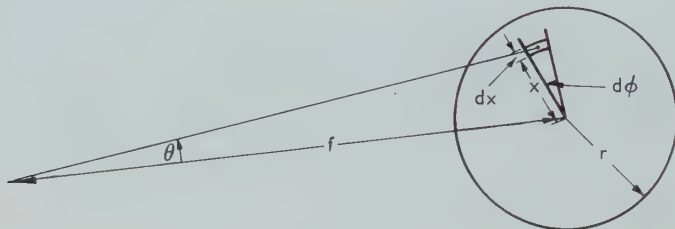


Fig. 1

The displacement of the fringe maxima and minima from their positions corresponding to an ideal point source will be determined in each case by finding the values of K for which $\partial I/\partial K = 0$. For the narrow slit, using differentiation under the integral sign,

$$\frac{\partial I_1}{\partial K} \propto \int_0^{\theta_r} \sin(2K \cos \theta) \cos \theta d\theta.$$

Putting $\cos \theta = 1 - \frac{1}{2}\theta^2$ we have

$$\frac{\partial I_1}{\partial K} \propto \sqrt{\left(\frac{\pi}{2K}\right)} \{ \sin 2KC(u) - \cos 2KS(u) \},$$

where

$$S(u) = \int_0^u \sin \frac{1}{2}\pi z^2 dz,$$

$$C(u) = \int_0^u \cos \frac{1}{2}\pi z^2 dz,$$

are the Fresnel integrals and

$$u = \sqrt{\left(\frac{2K}{\pi}\right)} \theta_r.$$

Thus we have $\partial I_1 / \partial K = 0$ when

$$K = \frac{1}{2} \tan^{-1} \frac{S(u)}{C(u)} + \frac{1}{2} n\pi \quad (n=0, 1, 2, 3, \dots), \quad \dots \quad (3)$$

whereas for two-beam fringes from an ideal point source $\partial I / \partial K = 0$ when $K = \frac{1}{2} n\pi$. With an accuracy of a few parts in a million, K may be taken as $2\pi l / \lambda$ in the arguments of the Fresnel integrals, l being the length of the end bar and the correction δ_1 to be applied is thus

$$\delta_1 = \frac{\lambda}{4\pi} \tan^{-1} \frac{S(v)}{C(v)}, \quad \dots \quad (4)$$

with

$$v = \sqrt{\left(\frac{4l}{\lambda}\right)} \theta_r.$$

The maximum value of the correction is approximately 0.17 fringe which occurs when the ratio of the Fresnel sine and cosine integrals is a maximum. This position corresponds to a phase difference $\frac{1}{2} K \theta_r^2$ between central and extreme rays through the aperture of the order of half a wavelength which corresponds to the Rayleigh criterion for two-beam fringes of good definition. For an end bar of length 300 mm, the fringe displacement for a narrow slit of length 0.3 mm used with a collimator of focal length 208 mm (which is a typical case occurring in practice) is obtained from equation (4) as 0.11 fringe. The displacement for a slit of length 0.6 mm is almost the same.

Observations showed that the displacement (if any) on increasing the slit length from 0.3 to 0.6 mm was verified by several observers to be definitely less than 0.1 fringe. Application of the correction formula $(a^2 + b^2) / 24f^2$ per unit length indicates that the fringe displacement would be approximately 0.3 fringe. The correction $2\delta_1 / \lambda$ (- fringe displacement) as a function of $\sqrt{(4l/\lambda)} \theta_r (= 2\sqrt{\Delta/\pi})$ is shown in Figure 2. Points on the axis where $\Delta = \pi/2, \pi, 2\pi, \dots$ are also shown.

A similar result to that given in equation (4) can be obtained for the circular aperture

$$\begin{aligned} \frac{\partial I_2}{\partial K} &\propto \int_0^{\theta_r} \sin(2K \cos \theta) \cos \theta d\theta \\ &= \theta_r I(\theta_r) - \int_0^{\theta_r} I(\theta) d\theta \end{aligned}$$

integrating by parts, where

$$I(\theta) \equiv \int_0^\theta \sin(2K \cos \theta) \cos \theta d\theta,$$

which was evaluated for the case of the slit and

$$I(\theta_r) = \int_0^{\theta_r} \sin(2K \cos \theta) \cos \theta d\theta.$$

Thus $\partial I_2 / \partial K = 0$ when

$$\theta_r I(\theta_r) = \int_0^{\theta_r} I(\theta) d\theta,$$

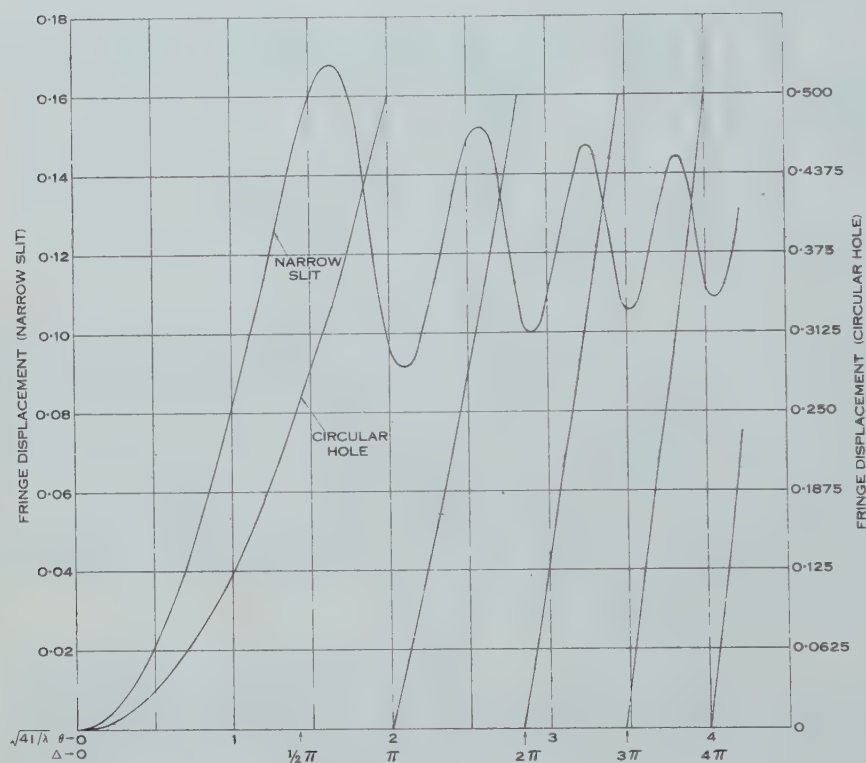


Fig. 2

which leads to a correction δ_2 of amount

$$\delta_2 = \frac{\lambda}{4\pi} \tan^{-1} \left[\frac{S(r) - \frac{1}{r} \int_0^r S(r) dr}{C(r) - \frac{1}{r} \int_0^r C(r) dr} \right]. \quad \dots \dots \dots (5)$$

A plot of δ_2 as a function of $\sqrt{(4l/\lambda)}\theta_r$ is given in Figure 2 also. Note that in this case the phase difference between central and extreme rays, $\frac{1}{2}K\theta_r^2$, is a multiple of π at positions of maximum fringe displacement. The curve obtained agrees precisely with that obtained in paper I by a different approach.

The correction δ_1 and δ_2 of course both reduce to zero for $\theta_r = 0$.

(b) *Multiple-Beam Systems*

In multiple-beam systems the effect of obliquity is not as important as for the two-beam case since the path differences involved are necessarily very small. The case of a rectangular aperture used in a multiple-beam interferometer will be discussed as an example. For the transmission case the intensity distribution is given by

$$T_4 = \int_{-a/2}^{a/2} \int_{-b/2}^{b/2} \frac{dx dy}{1 + R^2 \sin^2 (K \cos \theta)} \\ = \frac{4bf}{2 + R^2} \int_0^{\theta_r} \frac{d\theta}{1 - C \cos (2K \cos \theta)}, \quad \dots \dots \dots (6)$$

where $R^2 = 4r^2/(1-r^2)^2$ — Fabry's "Coefficient of Finesse" and $C = R^2/(2 + R^2)$. Expanding the integrand in equation (6) as a Fourier series with $2K \cos \theta$ as the independent variable we have

$$T_4 = \frac{2bf}{2 + R^2} \int_0^{\theta_r} [1 + \sum_{n=1}^{\infty} A_n \cos (2nK \cos \theta)] d\theta, \quad \dots \dots \dots (7)$$

where

$$A_n = \frac{2}{\pi} \int_0^{\pi} \frac{\cos nz}{1 - C \cos z} dz = \frac{2}{\sqrt{1 - C^2}} \left[\frac{1 - \sqrt{1 - C^2}}{C} \right]^n.$$

Therefore

$$\frac{\partial T_4}{\partial K} \propto \int_0^{\theta_r} \sum_{n=1}^{\infty} A_n n \cos \theta \sin (2nK \cos \theta) d\theta \\ = \sum_{n=1}^{\infty} A_n n \sqrt{\frac{\pi}{2nK}} [\sin 2nKC(\sqrt{n}u) - \cos 2nKS(\sqrt{n}u)].$$

The same form can be obtained for the reflection case.

Now multiple-beam interference conditions require K to be small ($K \sim 10^3$), so the argument of the Fresnel integrals, for θ_r equal to 0.001 and n as large as 40, is approximately 0.2 and up to such values the Fresnel sine integral is very small and may be neglected in comparison with the Fresnel cosine integral and the latter is very closely given by

$$\int_0^{\beta} \cos \frac{1}{2} \pi z^2 dz = \beta.$$

Therefore

$$\frac{\partial T_4}{\partial K} \propto \sum_{n=1}^{\infty} A_n n \sin 2nK, \\ \frac{\partial R_4}{\partial K} \propto \sum_{n=1}^{\infty} D_n n \sin 2nK,$$

and

$$\frac{\partial T_4}{\partial K} = \frac{\partial R_4}{\partial K} = 0,$$

when $K = 0, \frac{1}{2}\pi, \pi, \dots$ as for an ideal point source.

Thus it may be concluded that the correction due to the above effect is negligible for multiple-beam systems. The implied symmetry of the fringes in the above analysis is not correct, for consideration of the effect of phase changes introduced by multiply reflected beams shows that the fringes will be asymmetrical (Brossell 1947).

III. THE INTENSITY DISTRIBUTIONS

The fringe intensity distributions are themselves not of great importance but can be found if required in several ways in the different cases.

The formulae for the two-beam intensity distributions and those for the multiple-beam cases, can in each instance be shown to reduce to the corresponding formula for an ideal point source provided K is not too large. As an example the case of a rectangular aperture in a two-beam system will be examined.

Substituting $y = f \tan \theta$ in equation (1), we obtain

$$I_1 = K^2 b f \int_0^{\theta_r} \frac{\cos^2 (K \cos \theta)}{(K \cos \theta)^2} d\theta, \dots\dots\dots (8)$$

which can be compared with the expression $\int \sin^2 (K \sin \alpha) d\alpha / (K \sin \alpha)^2$ occurring in the formula for the diffracted energy from a rectangular aperture. The integral in equation (8) can be evaluated as follows using a Maclaurin expansion :

$$\cos^2 (K \cos \theta) = \cos^2 K + K \theta^2 \sin 2K - \frac{\theta^4}{4!} \{6K^2 \cos 2K + K \sin 2K\} + \dots$$

Therefore

$$I_1 = b f \left[\cos^2 K J_1 + K \sin 2K J_2 - \frac{(6K^2 \cos 2K + K \sin 2K)}{4!} J_3 + \dots \right],$$

where

$$\begin{aligned} J_1 &= \tan \theta_r, \\ J_2 &= \theta_r^2 \tan \theta_r + 2 \sum_{n=1}^{\infty} \frac{(-1)^n 2^{2n} (2^{2n} - 1)}{(2 + 2n - 1)(2n)!} B_{2n} \theta_r^{2+2n-1}, \\ J_3 &= \theta_r^4 \tan \theta_r + 4 \sum_{n=1}^{\infty} \frac{(-1)^n 2^{2n} (2^{2n} - 1)}{(4 + 2n - 1)(2n)!} B_{2n} \theta_r^{4+2n-1}, \end{aligned}$$

and B_{2n} are Bernoulli's numbers,

$$B_2 = -\frac{1}{30}, \quad B_4 = \frac{1}{30}, \quad B_6 = 0.253114, \quad B_8 = 7.092157, \dots$$

Taking $\tan \theta_r = \theta_r = a/f$, I_1 is given by

$$\begin{aligned} I_1 &= ab [\cos^2 K + 0.978 \theta_r^2 K \sin 2K - 0.00171 \theta_r^4 \{6K^2 \cos 2K + K \sin 2K\} + \dots] \\ &= ab \cos^2 K \quad \text{for } \theta_r \sim 0.001 \text{ and } K \gg 10^5 \text{ or so,} \end{aligned}$$

IV. CONCLUSIONS

The effect of the finite size of the aperture in practice does not exceed 0.17 fringe approximately for a narrow slit and 0.5 fringe for a circular aperture and can be conveniently found using the formulae and curves given in this paper.

The usual correction due to the finite size of the aperture appears to be invalid for long lengths.

V. ACKNOWLEDGMENT

The author expresses his appreciation of the many discussions on the subject of obliquity corrections with Mr. C. F. Bruce of the Division of Metrology.

VI. REFERENCES

- BROSSELL, J. (1947).—*Proc. Phys. Soc. Lond.* **59**: 224.
BRUCE, C. F. (1955).—*Aust. J. Phys.* **8**: 224.

THEORY OF TRACKS IN NUCLEAR RESEARCH EMULSIONS

By J. M. BLATT*

[Manuscript received October 19, 1954]

Summary

The process of the formation of a visible track in a nuclear research emulsion is approximated by a simplified model which reduces it to a one-dimensional problem. Within this basic approximation, there are included many different detailed models, including all models so far proposed in the literature. We then consider the following quantities characteristic of the track: the number of gaps longer than some minimum length r contained in a cell length t of the track, $n_r(t)$; and the sum of the excess (over and above r) lengths of these n_r gaps, $x_r(t)$. We derive complete theoretical results for the mean values and standard deviations of these quantities, in terms of a small number of definite integrals. We then specialize to particular detailed models, and make a quick comparison against experiment. The model of O'Ceallaigh fails to agree with experimental data, and so do all straightforward generalizations of that model, such as the model of Happ, Hull, and Morrish. The reason for this failure is traced to the neglect of the "graininess" of the emulsion before development. The model of Herz, which takes the extreme opposite point of view, comes much closer to fitting the data. We also give a reasonable modification of the Herz model. A more detailed comparison against experiment must still be made.

I. INTRODUCTION, DEFINITIONS, AND BASIS APPROXIMATIONS

A photographic emulsion consists of crystals of roughly spherical shape suspended in gelatin. In G5 emulsions, roughly half the volume is taken up by the crystals, the other half is gelatin. A charged particle passing through the emulsion activates some of the crystals through which it travels. When the emulsion is developed, each "activated" crystal grows into a "grain" of considerably larger radius (approximately twice the radius of the undeveloped crystal under usual conditions of development), while the inactive crystals dissolve away. If two activated crystals are close enough together, the final grains can overlap and form a "blob". Since the size of the final grains is close to the limit of optical resolution, it is generally difficult to tell under the microscope whether a blob consists of one, two, or three grains; a single large grain may be mistaken for a two-grain blob, for example. Hence it has become customary to do "grain-counting" by counting blobs (or, equivalently, by counting the gaps between the blobs). It is then necessary to develop a theoretical expression which relates the observed number of gaps per unit length of the track, or some other measurable quantity characteristic of the track, to the fundamental physical characteristics of the emulsion (initial crystal

*The F. B. S. Falkiner Nuclear Research and Adolph Bassar Computing Laboratories, School of Physics, University of Sydney; also supported by the Nuclear Research Foundation within the University of Sydney.

size, distribution of initial crystals in space, final grain size) and of the charged particle (probability of activation of a traversed crystal). Furthermore, since the process of track formation is statistical in character, we need to know the standard deviations which must be attached to the measured track characteristics. Indeed, the original impetus for this study arose from the experimental observation that the standard deviation of the number of gaps in a certain length of track l (the "cell length") is less than the standard deviation in a Poisson distribution, and furthermore seems to depend on the character of the track ("thin", "medium", or "thick"). It is the purpose of this paper to provide the necessary theoretical expressions to analyse experimental data.

The main difficulty in such a calculation lies in the spatial distribution of the initial grains. If we could assume that the initial grains traversed by the charged particle all have their centres along the path of the particle, the theory would be greatly simplified. We now present qualitative arguments to show that this is a very good approximation under normal conditions. Figure 1 (a)

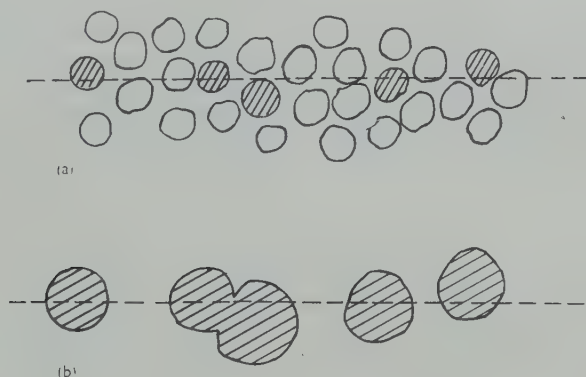


Fig. 1

gives a schematic picture of the emulsion as it may actually be, and indicates the path of the charged particle. Let us assume that some but not all of the crystals along this path are activated; the activated ones are indicated by shading. Figure 1 (b) shows this same track after development. Each activated crystal has grown to roughly twice the original size, the other crystals have dissolved. Some of the final grains have combined into blobs, others have remained single.

Figure 2 (a) shows the same track in the "linear approximation", in which we assume that the centre of each activated grain lies directly on the path of the charged particle (since the inactive grains are never seen, no assumption has to be made about their centres). Figures 1 (a) and 2 (a) differ appreciably. However, let us now look at the result after development. Figure 2 (b) shows the same track after development, again in the linear approximation. We see that the growth of the crystals into grains has had the result of decreasing the difference between Figures 1 (b) and 2 (b) compared to the difference between Figures 1 (a) and 2 (a). Indeed, the same grains which had previously merged

into blobs still do so in the linear approximation, except for a few borderline cases, perhaps. The larger the growth of the crystals during the process of development, the less important is the exact location of the centre of each grain. Hence we conclude that the linear approximation is adequate for this problem if there is appreciable growth of the crystals during development. Conversely, if it were possible to find a method of developing the emulsion in such a way that the activated crystals would not grow appreciably, the linear approximation would be of very doubtful validity.

For present methods of development, which result in a grain growth of roughly a factor of two, we feel that the linear approximation is entirely adequate. Indeed, a first approximation to the necessary correction terms can be obtained by observing that the random sideways displacements of the centres of activated

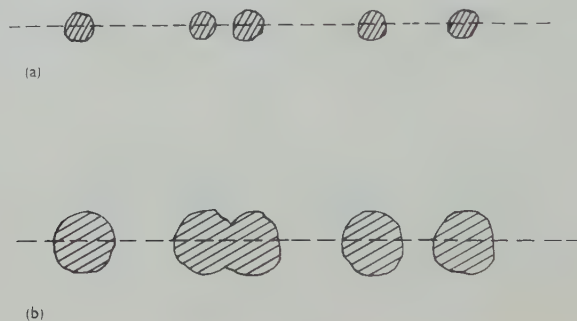


Fig. 2

crystals from the nearest point along the path of the charged particle are roughly equivalent to random changes in the final size of the grown grains. Since there is a distribution of final grain sizes in any case, the linear approximation is even more accurate than the comparison between Figures 1 and 2 indicates. From now on we shall adopt the linear approximation throughout this paper.

In this approximation, the visible characteristics of the track depend upon two probability distributions only :

- (1) The distribution-in-final-size of the grains after development. We let γ stand for the *radius* of the developed grain, and we let $G(\gamma)d\gamma$ be the probability that this radius lies between γ and $\gamma+d\gamma$.
- (2) The "survival probability" $Q(t)$; this is the probability that the distance between the centre points of adjacent activated grains exceeds t . Alternatively, $Q(t)$ is the probability that the charged particle, after activating a crystal along its path, "survives" for a distance larger than t before activating the next crystal.

We shall solve the statistical problems for two quantities of interest in the experimental study of tracks, for completely arbitrary functions $G(\gamma)$ and $Q(t)$. By specializing to various particular functions $G(\gamma)$ and $Q(t)$, we can then deduce results for particular models of track formation, and compare them with experiment.

Since the distribution in final grain size $G(\gamma)$ is experimentally a very sharp distribution, it is not necessary to explore all possible distribution functions. We shall give results for a delta function (infinitely sharp) $G(\gamma)$, and for a particular family of functions $G(\gamma)$ which is chosen so that (a) the integrals can be evaluated in closed form, (b) there are two adjustable parameters in $G(\gamma)$ which can be used to fit the mode and the standard deviation of the observed distribution in sizes of single grains (as seen, for example, by studying a very thin track).

The simplest assumption we can make about the survival probability $Q(t)$ is the one made by O'Ceallaigh (1953) as well as by Happ, Hull, and Morrish (1952). These models differ only in the distribution $G(\gamma)$, which is assumed to be infinitely sharp by O'Ceallaigh but not by Happ, Hull, and Morrish. The survival probability in both models is the same, and is obtained by assuming that there is a constant probability λdt per path length dt for activating the "next" crystal. The resulting survival probability is

$$Q(t) = e^{-\lambda t}, \quad \text{"model A"}. \quad \dots\dots\dots (1.1)$$

This model neglects the correlations between positions of crystal centres in the emulsion. It is obviously a good approximation if the emulsion is very dilute, or (equivalently) if the probability of the charged particle activating any crystal it traverses is very low. In either case the final appearance of the track is that of a "thin" track, with wide spaces separating the individual grains and only little blob formation.

The extreme opposite model was chosen by Herz (Herz and Davis 1955), who assumes that the crystals are arrayed along the path of the charged particle in a regular lattice of definite lattice spacing. The charged particle has a probability p for activating each crystal it traverses. Hence the survival probability $Q(t)$ is a step function:

$$\left. \begin{aligned} Q(t) &= 1, & 0 < t < \beta, \\ &= 1-p, & \beta < t < 2\beta, \\ &= (1-p)^2, & 2\beta < t < 3\beta, \\ &\dots\dots\dots \end{aligned} \right\} \text{"model B"}. \quad \dots\dots (1.2)$$

The true situation is surely somewhere between these two extreme assumptions. If each crystal is a sphere of diameter α , then clearly $Q(t)=1$ for $0 < t < \alpha$, in contradiction to model A. On the other hand, the emulsion is essentially amorphous, not crystalline, hence there can not be any long-range order. At distances of several crystal diameters, the "memory" of the location of the initial crystal is lost, and the survival probability becomes asymptotically equal to (1.1), in contradiction to model B. A schematic picture of $Q(t)$ for models A and B and for the true emulsion is given in Figure 3.

It might be thought that the main error in model A consists of the neglect of the size of the initial crystal; hence we have also calculated various quantities measured on tracks with a third model, which we shall call model C:

$$\left. \begin{aligned} Q(t) &= 1, & 0 < t < \alpha, \\ &= \exp[-\lambda(t-\alpha)], & t > \alpha. \end{aligned} \right\} \text{"model C"}. \quad \dots\dots (1.3)$$

We now define the "number of gaps longer than r ", $n_r(t)$, in a track of length t . We are free to define this quantity in such a way as to make the subsequent mathematics easy, and we shall do so. It should be emphasized that *our definition of a "countable gap", while simple enough to learn, is not identical*

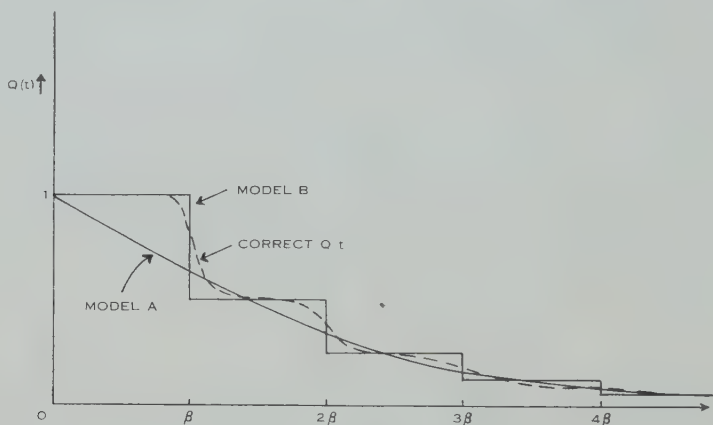


Fig. 3

with the usual one. The "last" gap is sometimes included in the count, sometimes excluded. This makes little difference if the cell length t is long enough, but it is important when comparing the theory against experiments with a short cell length. Since the experiments on the variance in the number of gaps necessarily use rather short cell lengths, this caution is important.

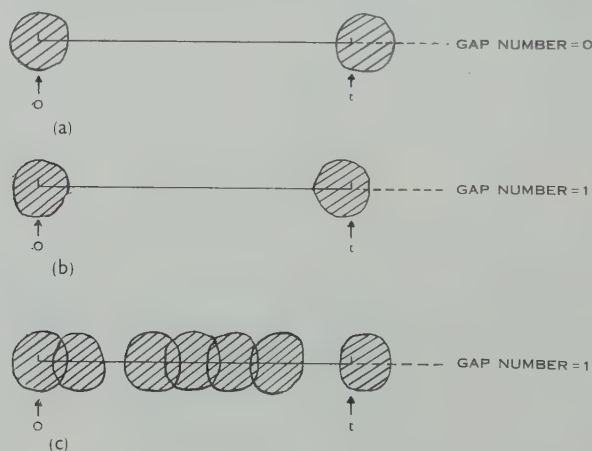


Fig. 4

Our definition of a "countable gap" consists of two parts:

- (1) The initial point $t=0$ of each cell length must be inside a blob.
- (2) The grain terminating each "countable" gap must have its *centre* point within the cell length t .

Hence, for example, the situation depicted in Figure 4 (a) corresponds to no gap at all. In general, our convention will result in omitting the "last" gap in thin tracks, while counting it in thick tracks.

The quantity $n_r(t)$ is defined as the number of countable gaps, each one of length equal to or larger than r , in a cell length t . The quantity $x_r(t)$ is the total length of these n_r gaps, counting for each gap only the *excess* of its length over r .

In Section II we give formal solutions, in terms of definite integrals, for the distribution-in- n_r and the distribution-in- x_r . This section is mathematical in character and can be omitted if desired. Sections III and IV contain the results in usable form, as well as a short discussion of the differences between the models and a short comparison with experimental data. An extended experimental check is now under way in this Laboratory, and will be the subject of a future publication.

II. A FORMAL SOLUTION USING THE REGENERATION POINT METHOD

Let $\pi_n(r, t)$ be the probability that the number $n_r(t)$ of countable gaps longer than r in a cell length t is equal to the integer n . We now derive an equation for $\pi_n(r, t)$ by using the "regeneration point method" of Bellman and Harris (1948) and Janossy (1950). We concentrate our attention on the position t' of the "next" activated crystal. The "position" of any crystal is that of its centre point, by definition, no matter which point within the crystal was actually activated. The assumption is that during development each activated crystal grows as a whole, so that the precise point of activation within the crystal does not matter.

The probability $q(t')dt'$ that the centre of the "next" activated crystal lies between t' and $t'+dt'$ is simply related to the survival probability $Q(t')$; namely,

$$q(t') = -dQ/dt'. \quad \dots\dots\dots (2.1)$$

We shall use (2.1) formally even when the process of differentiation is not permissible mathematically (e.g. when $Q(t)$ is the step function of model B). A rigorous mathematical treatment can be given in terms of Stieltje's integrals, but the final result is the same.

Depending upon the position t' of the next activated crystal, and upon the final radii γ_1 and γ_2 of the initial and the "next" grains, there are three possibilities to consider:

- (1) The grains growing from the initial and the next crystal merge to form a blob, or are so close that the resultant gap is of length smaller than our minimum length r . This happens when $t' < \gamma_1 + \gamma_2 + r$. In this case the remaining distance, $t - t'$, must contain n gaps of length greater than r , in order to have a contribution to $\pi_n(r, t)$. This contribution is therefore

$$\int_0^\infty d\gamma_1 \int_0^\infty d\gamma_2 \int_0^{\gamma_1 + \gamma_2 + r} dt' q(t') G(\gamma_1) G(\gamma_2) \pi_n(r, t - t').$$

- (2) The "next" crystal is far enough away to form a gap of length greater than r , but its centre is still within our cell length, that is, we have $\gamma_1 + \gamma_2 + r < t' < t$. The remaining distance, $t - t'$, must now contain $n - 1$ gaps of length greater than r , in order to contribute to $\pi_n(r, t)$. The contribution to $\pi_n(r, t)$ from this range of values of t' is

$$\int_0^\infty d\gamma_1 \int_0^\infty d\gamma_2 \int_{\gamma_1 + \gamma_2 + r}^t dt' q(t') G(\gamma_1) G(\gamma_2) \pi_{n-1}(r, t - t').$$

- (3) The final possibility is that the next activated crystal lies outside the cell length t altogether, that is, $t' > t$. According to our conventions for "countable" gaps, we then say that there is no gap at all in our cell length. Hence possibility (3) makes a contribution to $\pi_0(r, t)$ only. We define the quantity δ_n to equal unity when $n = 0$, to equal zero otherwise. Then the contribution of this range of t' to $\pi_n(r, t)$ is

$$\int_t^\infty dt' q(t') \delta_n = Q(t) \delta_n.$$

We now observe that the possibilities (1), (2), and (3) above correspond to mutually exclusive events, of which exactly one must happen. Hence we can add the probabilities in order to get $\pi_n(r, t)$, giving the regeneration point equation

$$\begin{aligned} \pi_n(r, t) = & \int_0^\infty d\gamma_1 \int_0^\infty d\gamma_2 \int_0^{\gamma_1 + \gamma_2 + r} dt' q(t') G(\gamma_1) G(\gamma_2) \pi_n(r, t - t') \\ & + \int_0^\infty d\gamma_1 \int_0^\infty d\gamma_2 \int_{\gamma_1 + \gamma_2 + r}^t dt' q(t') G(\gamma_1) G(\gamma_2) \pi_{n-1}(r, t - t') \\ & + Q(t) \delta_n. \end{aligned} \quad (2.2)$$

The actual ranges of integration over t' are restricted somewhat more than indicated in (2.2) by the condition*

$$\pi_n(r, t) = 0, \quad \text{unless } t > nr. \quad (2.3)$$

We now solve (2.2) by means of Laplace transforms. We introduce

$$p_n(r, s) = \int_0^\infty e^{-st} \pi_n(r, t) dt. \quad (2.4)$$

We multiply both sides of (2.2) by e^{-st} and integrate over t . We then interchange the orders of integration with respect to t and t' on the right side. It is useful to introduce the notations:

$$h(s) = s \int_0^\infty e^{-st} Q(t) dt = 1 - \int_0^\infty e^{-st} q(t) dt, \quad (2.5)$$

$$g(r, s) = \int_0^\infty d\gamma_1 \int_0^\infty d\gamma_2 \int_{\gamma_1 + \gamma_2 + r}^\infty dt e^{-st} q(t) G(\gamma_1) G(\gamma_2). \quad (2.6)$$

* This condition can be made sharper if the final grain radii γ cannot be smaller than a minimum value γ_{\min} , i.e. if $G(\gamma) = 0$ for $\gamma < \gamma_{\min}$. Then the condition on $\pi_n(r, t)$ is: $t > n(r + 2\gamma_{\min})$.

We then get the following Laplace transform equation from (2.2):

$$p_n(r, s) = (1 - h - g)p_n + gp_{n-1} + s^{-1}h(s)\delta_n, \quad \dots\dots\dots (2.7)$$

We set $n=0$ to get

$$p_0(r, s) = \frac{h(s)}{s[h(s) + g(r, s)]}, \quad \dots\dots\dots (2.8)$$

We introduce the notation

$$K(r, s) = \frac{g(r, s)}{h(s) + g(r, s)}, \quad \dots\dots\dots (2.9)$$

and obtain the *complete formal solution*

$$p_n(r, s) = [K(r, s)]^n p_0(r, s). \quad \dots\dots\dots (2.10)$$

The equation (2.10) together with the defining equations (2.5), (2.6), (2.8), and (2.9) constitutes a formal solution to the statistical problem of the number of gaps larger than r . The problem has been reduced to the computation of certain definite integrals, $h(s)$ and $g(r, s)$, and the subsequent inversion of the Laplace transforms (2.10). In practice the integrals $h(s)$ and $g(r, s)$ present little difficulty, but the inversion of the Laplace transform is usually not possible in closed form in terms of tabulated functions.

In practice, experimentalists are not interested in the probability $\pi_n(r, t)$ of getting exactly n gaps of length in excess of r . They would like to know the mean number of these gaps, and the standard deviation around this mean. These quantities are related to the moments of the distribution in $n_r(t)$. We introduce the factorial moments $F_k(r, t)$ by

$$F_k(r, t) = \langle n_r(n_r - 1)(n_r - 2) \dots (n_r - k + 1) \rangle_{\text{Ave}} = \sum_{n=0}^{\infty} n(n-1) \dots (n-k+1) \pi_n(r, t). \quad \dots\dots\dots (2.11)$$

The first few factorial moments are:

$$F_0(r, t) = 1, \quad \dots\dots\dots (2.12)$$

$$F_1(r, t) = \langle n_r(t) \rangle_{\text{Ave}}, \quad \dots\dots\dots (2.13)$$

$$F_2(r, t) = \langle n_r^2(t) \rangle_{\text{Ave}} - \langle n_r(t) \rangle_{\text{Ave}}^2, \quad \dots\dots\dots (2.14)$$

and the standard deviation is given by

$$\sigma^2 = \overline{n^2} - (\bar{n})^2 = F_2 - F_1^2 + F_1. \quad \dots\dots\dots (2.15)$$

The Laplace transforms of the factorial moments are given by

$$f_k(r, s) = \int_0^{\infty} e^{-st} F_k(r, t) dt = \sum_{n=0}^{\infty} n(n-1) \dots (n-k+1) p_n(r, s), \quad \dots (2.16)$$

and substitution of (2.10) into (2.16) gives

$$f_k(r, s) = k! p_0(r, s) \frac{[K(r, s)]^k}{[1 - K(r, s)]^{k+1}} \\ = \frac{k!}{s} \left[\frac{g(r, s)}{h(s)} \right]^k \dots \dots \dots (2.17)$$

The Laplace transform (2.17) has a particularly simple structure, as a result of certain properties of the functions $h(s)$ and $g(r, s)$ which we shall now point out. We see from the definition (2.5) of $h(s)$ that $h(s)=0$ when $s=0$: hence (2.17) has a pole of order $k+1$ at the origin in the s -plane. Next we observe that $|h(s)| > 0$ as long as the real part of s is greater than zero. Also $h(s^*) = [h(s)]^*$ so that all its zeros must occur in complex conjugate pairs. The function $g(r, s)$ defined by (2.6) has no poles in the positive half-plane of s , nor at $s=0$. g may have poles along the negative s -axis, or pairs of complex conjugate poles in the negative half-plane of s . We therefore conclude that $f_k(r, s)$ has a pole of order $k+1$ at $s=0$, no poles in the positive half-plane of s , and (perhaps) pairs of conjugate complex poles in the negative half-plane of s .

When the Laplace transform (2.17) is inverted by the method of residues, a pole at a point s_1 contributes a term proportional to $e^{s_1 t}$, unless $s_1=0$, in which case we get a polynomial in t . Since all poles except the one at $s=0$ lie in the negative half-plane of s , their contributions are damped exponentials in t : hence if t is large enough, these contributions can all be neglected.

We therefore restrict ourselves to the contribution of the pole at $s=0$, and obtain

$$F_k(r, t) = \frac{d^k}{ds^k} \left\{ e^{st} \left[\frac{sg(r, s)}{h(s)} \right]^k \right\}_{s=0}, \\ F_k(r, t) = \sum_{k'=0}^k \frac{k!}{k'!(k-k')!} t^{k-k'} \left\{ \frac{d^{k'}}{ds^{k'}} \left[\frac{sg(r, s)}{h(s)} \right]^k \right\}_{s=0} \dots (2.18)$$

Thus the problem of finding the moments of the distribution-in- $n_r(t)$ has been reduced to the evaluation of the definite integrals (2.5) and (2.6) and simple differentiations. We emphasize that this has been achieved without any major loss of generality: the distribution $G(\gamma)$ of final grain radii and the survival probability $Q(t)$ are still completely arbitrary and at our disposal, subject only to the (not very stringent) condition that the integrals (2.5) and (2.6) can be done in closed form.

We now turn our attention to the distribution in the total gap length $x_r(t)$. $x_r(t)$ is defined as the sum of the "excess" lengths of all countable gaps longer than r . We emphasize that gaps shorter than r make no contribution to $x_r(t)$, while gaps longer than r (provided they are "countable" by our conventions) contribute not their whole length, but merely the excess of their length over r . We now let $\pi(r, x, t)dx$ be the probability that the total gap length $x_r(t)$ lies in the range $x, x+dx$. Since there exists a non-zero probability of having no gaps at all of length greater than r , $\pi(r, x, t)$ contains a singular contribution equal to $\pi_0(r, t)\delta(x)$, where $\pi_0(r, t)$ is the probability of finding no gaps at all (longer

than r) in the cell length t , and $\delta(x)$ is Dirac's delta function. We shall show the existence of this contribution in our solution later on.

By arguments entirely analogous to the ones used in deriving equation (2.2), we find the following regeneration point equation for $\pi(r, x, t)$:

$$\begin{aligned} \pi(r, x, t) = & \int_0^\infty d\gamma_1 \int_0^\infty d\gamma_2 \int_0^{\gamma_1 + \gamma_2 + r} dt' q(t') G(\gamma_1) G(\gamma_2) \pi(r, x, t - t') \\ & + \int_0^\infty d\gamma_1 \int_0^\infty d\gamma_2 \int_{\gamma_1 + \gamma_2 + r}^t dt' q(t') G(\gamma_1) G(\gamma_2) \pi(r, x - t' + \gamma_1 + \gamma_2 + r, t - t') \\ & + Q(t) \delta(x). \end{aligned} \quad (2.19)$$

We solve this equation by Laplace transforms; this time we need a double Laplace transform, on the variable x as well as on the variable t . We define

$$\varphi(r, u, s) = \int_0^\infty dx \int_0^\infty dt \exp(-ux - st) \pi(r, x, t). \quad (2.20)$$

We multiply both sides of (2.19) by $\exp(-ux - st)$ and integrate over x and t . In addition to the functions $h(s)$ and $g(r, s)$ encountered previously, we need to define an additional function

$$\begin{aligned} v(r, u, s) = & \int_0^\infty d\gamma_1 \int_0^\infty d\gamma_2 \int_{\gamma_1 + \gamma_2 + r}^\infty dt q(t) G(\gamma_1) G(\gamma_2) \\ & \times \exp[(\gamma_1 + \gamma_2 + r)u - (u + s)t]. \end{aligned} \quad (2.21)$$

We observe that

$$v(r, 0, s) = g(r, s), \quad (2.22)$$

and

$$\lim_{u \rightarrow \infty} v(r, u, s) = 0. \quad (2.23)$$

The Laplace transform image of equation (2.19) then is

$$\varphi(r, u, s) = [1 - h(s) - g(r, s)] \varphi(r, u, s) + v(r, u, s) \varphi(r, u, s) + s^{-1} h(s),$$

which gives the formal solution

$$\varphi(r, u, s) = \frac{h(s)}{s[h(s) + g(r, s) - v(r, u, s)]}. \quad (2.24)$$

In the general case, the function $v(r, u, s)$ may be difficult to find in closed form, and the subsequent inversion of the double Laplace transform (2.24) can be very difficult indeed. Hence (2.24) as it stands is not directly a practical solution. However, we shall show that information of practical value can be obtained by using (2.24).

First of all, we perform two checks on (2.24):

- (1) When $u=0$, $\varphi(r, 0, s)$ must be the Laplace transform of the probability of getting any gap length x whatever, by comparison with (2.20); substitution of (2.22) into (2.24) gives $\varphi(r, 0, s) = s^{-1}$, which is indeed the Laplace transform of unity.

- (2) $\varphi(r, u, s)$ must contain the Laplace transform of the singular contribution $\pi_0(r, t)\delta(x)$, that is, the quantity $p_0(r, s)$, equation (2.8). We can isolate this contribution by going to the limit of large positive u , since all non-singular functions of x have Laplace transforms which vanish for large positive u . Substitution of (2.23) into (2.24) shows that $\lim_{u \rightarrow \infty} \varphi(r, u, s) = p_0(r, s)$, as it should be.

Next, we observe that we are primarily interested in the moments of the distribution in $x_r(t)$. Multiplication of $\pi(r, x, t)$ by x^k and integration over x has a simple Laplace transform image: k -fold differentiation with respect to the transform variable u , multiplication by $(-1)^k$, and setting $u=0$. We therefore define the quantity $\varphi_k(r, s)$ by

$$\varphi_k(r, s) = (-1)^k \left[\frac{\partial^k}{\partial u^k} \varphi(r, u, s) \right]_{u=0}. \quad (2.25)$$

Then $\varphi_1(r, s)$ is the Laplace transform of $\overline{x_r(t)}$, $\varphi_2(r, s)$ is the Laplace transform of $\overline{x_r^2(t)}$, and so on. Since higher moments give rather complicated expressions, we shall restrict ourselves to the first two moments, which are sufficient to give the mean value and the standard deviation.

In differentiating (2.24) with respect to u , we notice that the only quantity depending on u is $v(r, u, s)$. It helps to introduce a notation for the derivatives of v . We define

$$\begin{aligned} g_k(r, s) &= (-1)^k \left[\frac{\partial^k v(r, u, s)}{\partial u^k} \right]_{u=0} \\ &= \int_0^\infty d\gamma_1 \int_0^\infty d\gamma_2 \int_{\gamma_1 + \gamma_2 + r}^\infty dt (t - \gamma_1 - \gamma_2 - r)^k q(t) G(\gamma_1) G(\gamma_2) e^{-st}. \end{aligned} \quad (2.26)$$

Clearly we have

$$g_0(r, s) = g(r, s). \quad (2.27)$$

We then find from (2.24) and (2.25), using (2.22),

$$\varphi_1(r, s) = \frac{g_1(r, s)}{sh(s)}, \quad (2.28)$$

and

$$\varphi_2(r, s) = \frac{g_2(r, s)}{sh(s)} + \frac{2}{s} \left[\frac{g_1(r, s)}{h(s)} \right]^2. \quad (2.29)$$

We now observe that these Laplace transforms again have poles at $s=0$ and no poles in the positive half-plane of s . We can again restrict ourselves to the contribution from the pole at $s=0$. This contribution is always a polynomial in the cell length t . We introduce some more notation, to simplify writing,

$$w_{kk'}(r) = \frac{d^k}{ds^k} \left[\frac{sg_{k'}(r, s)}{h(s)} \right]_{s=0}. \quad (2.30)$$

We then get the following explicit expressions :

$$\overline{x_r(t)} = w_{01}t + w_{11}, \quad (2.31)$$

$$\overline{x_r^2(t)} = (w_{01})^2 t^2 + (4w_{01}w_{11} + w_{02})t + [2(w_{11})^2 + 2w_{01}w_{21} + w_{12}]. \quad (2.32)$$

A quantity of more interest than the mean square of $x_r(t)$ is the variance, which is given by

$$\overline{x_r^2(t)} - (\overline{x_r(t)})^2 = (2w_{01}w_{11} + w_{02})t + [(w_{11})^2 + 2w_{01}w_{21} + w_{12}]. \quad \dots (2.33)$$

We now continue the reduction of these expressions to practical form. We have already neglected the contributions from poles other than the dominant pole at $s=0$. We now make an additional approximation: since the cell length t is in general much larger than the average final grain radius $\bar{\gamma}$, the constant terms in (2.31) and (2.33), as well as in corresponding expressions for the mean value and variance of the number of gaps, are small corrections to the terms proportional to t . It turns out that the terms proportional to t involve fewer differentiations, and are therefore easier to evaluate practically, than the constant terms. On the other hand, the differentiations necessary to find the coefficient of t in the variance already suffice to determine also the constant term in the mean value. Hence we shall make the following approximation: we give the mean values exactly (i.e. neglecting only the contributions of other poles, which we shall show to be an excellent approximation), but we shall give only the coefficient of t (i.e. we shall neglect the constant term) for the variance. Thus the expressions for the variance will be somewhat less accurate than for the mean values; this is sufficient for comparison with experiment, since the variance is much harder to determine accurately anyway. In the limit of very large cell length t the expression for the variance is accurate; in practice, the fractional error is of order $(2\bar{\gamma}+r)/t$, where $\bar{\gamma}$ is the mean radius of developed grains, r is the minimum gap length chosen for counting, and t is the cell size. Thus, in G5 emulsions, we have $\bar{\gamma}=0.3 \mu$ and the error is less than 2 per cent. if $r=0$ and the cell length t exceeds 30μ (as it always does). An error of 2 per cent. in the variance is generally of no significance.

Making these approximations, it turns out that all the information of practical value can be reduced to the evaluation of the following seven definite integrals:

$$Q_0 = \int_0^\infty Q(t) dt, \quad \dots (2.34)$$

$$Q_1 = \int_0^\infty t Q(t) dt, \quad \dots (2.35)$$

$$g_{00} = \int_0^\infty d\gamma_1 \int_0^\infty d\gamma_2 \int_{\gamma_1+\gamma_2+r}^\infty dt q(t) G(\gamma_1) G(\gamma_2), \quad \dots (2.36)$$

$$g_{01} = \int_0^\infty d\gamma_1 \int_0^\infty d\gamma_2 \int_{\gamma_1+\gamma_2+r}^\infty dt t q(t) G(\gamma_1) G(\gamma_2), \quad \dots (2.37)$$

$$g_{10} = \int_0^\infty d\gamma_1 \int_0^\infty d\gamma_2 \int_{\gamma_1+\gamma_2+r}^\infty dt (t - \gamma_1 - \gamma_2 - r) q(t) G(\gamma_1) G(\gamma_2), \quad \dots (2.38)$$

$$g_{11} = \int_0^\infty d\gamma_1 \int_0^\infty d\gamma_2 \int_{\gamma_1+\gamma_2+r}^\infty dt t(t - \gamma_1 - \gamma_2 - r) q(t) G(\gamma_1) G(\gamma_2), \quad \dots (2.39)$$

$$g_{20} = \int_0^\infty d\gamma_1 \int_0^\infty d\gamma_2 \int_{\gamma_1+\gamma_2+r}^\infty dt (t - \gamma_1 - \gamma_2 - r)^2 q(t) G(\gamma_1) G(\gamma_2). \quad \dots (2.40)$$

In practice these integrals are evaluated most easily by Laplace transform methods; details are given in Appendix I. In terms of these integrals, the final results for the mean values are:

$$\overline{n_r(t)} = \frac{g_{00}t}{Q_0} + \frac{g_{00}Q_1 - g_{01}Q_0}{(Q_0)^2}, \quad \dots\dots\dots (2.41)$$

$$\overline{x_r(t)} = \frac{g_{10}t}{Q_0} + \frac{g_{10}Q_1 - g_{11}Q_0}{(Q_0)^2}, \quad \dots\dots\dots (2.42)$$

We shall describe the fluctuations around these mean values by the *fluctuation parameters* μ_n and μ_x which are defined to be the ratios of the standard deviations to the square roots of the mean values; for a Poisson distribution $\mu_n = \mu_x = 1$. We then get, restricting ourselves to the coefficients of t :

$$(\mu_n)^2 = \frac{(\overline{n_r^2}) - (\overline{n_r})^2}{\overline{n_r}} = 1 + 2 \frac{g_{00}Q_1 - g_{01}Q_0}{(Q_0)^2} + \text{order} \left(\frac{2\overline{\gamma} + r}{t} \right), \quad \dots\dots\dots (2.43)$$

$$(\mu_x)^2 = \frac{(\overline{x_r^2}) - (\overline{x_r})^2}{\overline{x_r}} = \left(\frac{g_{20}}{g_{10}} + 2 \frac{g_{10}Q_1 - g_{11}Q_0}{(Q_0)^2} \right) \left(1 + \text{order} \left(\frac{2\overline{\gamma} + r}{t} \right) \right), \quad \dots\dots (2.44)$$

The main error made in these expressions, indeed the only error in (2.41) and (2.42), is the neglect of the contribution of the other poles in the complex s -plane. This error becomes small (as we shall show in Section III) if there are many developed grains in the cell length t , on the average, so that t contains many "mean free paths" for grain activation. If we set $r=0$, the quantity $x_r(t) = \overline{x_0(t)}$ is the expectation value of the total length of the gaps; hence $t - \overline{x_0(t)}$ is the expectation value of the total length occupied by the blobs. If this length is much larger than a mean grain diameter $2\overline{\gamma}$, there are many developed grains (on the average) within the cell length t . Hence our *criterion of validity* reads

$$t - \overline{x_0(t)} \gg 2\overline{\gamma}, \quad \dots\dots\dots (2.45)$$

or, to the extent that we can neglect the constant term in (2.42),

$$t \gg \frac{2\overline{\gamma}Q_0}{Q_0 - g_{10}}, \quad \dots\dots\dots (2.46)$$

where g_{10} is evaluated for $r=0$.

III. RESULTS FOR SPECIFIC MODELS

(a) The Model of O'Ceallaigh

This model is characterized by the survival probability (1.1) together with the assumption of a unique value γ of the radius of the developed grains, that is, the distribution function $G(\gamma)$ is a delta function. We shall also use this model to test the validity of the approximations made in Section II, in particular the sufficiency of condition (2.46) on the minimum cell length required. This can

be done because the Laplace transforms can be inverted exactly. The exact results are

$$\begin{aligned}\overline{n_r(t)} &= 0, & \text{for } t < 2\gamma + r, \\ &= \lambda(t - 2\gamma - r) \exp[-\lambda(2\gamma + r)], & \text{for } t > 2\gamma + r, \\ & \dots\dots\dots (3.1)\end{aligned}$$

$$\begin{aligned}\overline{x_r(t)} &= 0, & \text{for } t < 2\gamma + r, \\ &= (t - 2\gamma - r - \lambda^{-1}) \exp[-\lambda(2\gamma + r)] + \lambda^{-1} \exp(-\lambda t), & \text{for } t > 2\gamma + r, \\ & \dots\dots\dots (3.2)\end{aligned}$$

whereas the results obtained by the use of (2.41) and (2.42) are

$$\overline{n_r(t)} = \lambda(t - 2\gamma - r) \exp[-\lambda(2\gamma + r)], \quad \dots\dots\dots (3.1a)$$

$$\overline{x_r(t)} = (t - 2\gamma - r - \lambda^{-1}) \exp[-\lambda(2\gamma + r)]. \quad \dots\dots (3.2a)$$

The difference between (3.1) and (3.1a) occurs only for $t < 2\gamma + r$; actually there is no real difference, we have simply not been careful in the general treatment about specifying the smallest value of t for which the pole at $s = 0$ makes a contribution. As we see from (3.1) this is not a significant restriction for any reasonable cell length t . The difference between (3.2) and (3.2a), for $t > 2\gamma + r$, is due to the contribution of a pole at $s = -1$. Condition (2.46) (or (2.45), from which it is derived) becomes

$$t \gg \frac{2\gamma}{1 - e^{-2\gamma\lambda}}. \quad \dots\dots\dots (3.3)$$

We now assert that the term $\lambda^{-1}e^{-\lambda t}$ in (3.2) is indeed negligible whenever (3.3) is satisfied. To see this, we take the extreme cases of a very thin and a very thick track. A thin track has $2\gamma\lambda \ll 1$, hence (3.3) becomes $t \gg \lambda^{-1}$, $\exp(-\lambda t) \ll 1$. A very thick track has $2\gamma\lambda \gg 1$, hence (3.3) is equivalent to $t \gg 2\gamma$, and thus $e^{-\lambda t} \ll e^{-2\gamma\lambda} \ll 1$, which establishes what we want. We also observe that condition (2.45) becomes trivial for thick tracks, and is always satisfied in practice. Even for thin tracks, (2.45) merely asserts that the cell length t must contain many developed grains, and hence many gaps. A cell length containing as few as five gaps on the average is sufficient to make the $\lambda^{-1}e^{-\lambda t}$ term in (3.2) less than 1 per cent. of the main term. Since no experimentalist does track statistics with such short cell lengths, condition (2.45) is in no sense a practical restriction.

The fluctuation parameters in the O'Ceallaigh model are given by

$$\mu_n^2 = 1 - 2\lambda(2\gamma + r) \exp[-\lambda(2\gamma + r)] + \text{order}[(2\gamma + r)/t], \quad \dots\dots\dots (3.4)$$

$$\mu_x^2 = (2/\lambda)\{1 - [1 + \lambda(2\gamma + r)] \exp[-\lambda(2\gamma + r)] + \text{order}[(2\gamma + r)/t]\}. \quad \dots (3.5)$$

For the purpose of testing the predictions of this model against experiment, the following relation (which follows from (3.1a) and (3.2a)) is very useful:

$$\overline{x_r(t)}/\overline{n_r(t)} = \lambda^{-1}[1 - \text{order}(1/\lambda t)]. \quad \dots\dots\dots (3.6)$$

For all practical cell lengths $\lambda t \gg 1$, so that (3.6) can be used to compute λ directly. Since all the subsequent formulas depend on the value of λ , this is a convenient procedure.

The ratio on the left side of (3.6) is approximately equal to the mean length of individual gaps.* In the O'Ceallaigh model, the parameter λ can assume any value; according to (3.1) and (3.2), very thick tracks are described by very large values of λ . We see from (3.6) that *the mean gap width approaches zero for very thick tracks*, on the O'Ceallaigh model. Indeed, we shall see that this is true for all "continuum models", that is, all models which ignore the "graininess" of the emulsion before development.

(b) *The Model of Happ, Hull, and Morrish*

This model has the same survival probability, (1.1), as the O'Ceallaigh model, but introduces a distribution $G(\gamma)$ of final grain radii γ . Rather than adopting the particular distribution function $G(\gamma)$ used by Happ, Hull, and Morrish, we shall work with a family of functions $G(\gamma)$ for which the integrals can be done in closed form. Since the experimentally observed distribution in γ is very sharp, the detailed shape of $G(\gamma)$ is of no great importance. We merely must allow ourselves two parameters in the functional form of $G(\gamma)$ which can be used to fit the mean value $\bar{\gamma}$ and the spread in the distribution-in- γ . We shall adopt the functional form

$$G(\gamma) = (k!)^{-1} (k/\gamma_0)^{k+1} \gamma^k \exp(-k\gamma/\gamma_0). \quad \dots\dots\dots (3.7)$$

For sharp distributions, that is, high values of k , γ_0 is close to the mean value $\bar{\gamma}$. More precisely,

$$\gamma_0 = \bar{\gamma} - \frac{\bar{\gamma}^2 - (\bar{\gamma})^2}{\bar{\gamma}}, \quad k = \frac{\gamma_0}{\bar{\gamma} - \gamma_0}. \quad \dots\dots\dots (3.8)$$

For typical experimental distributions of radii of single grains (*not* lengths of blobs which may contain several grains) k is in the neighbourhood of 10.

The quantities of experimental interest are then given by:

$$\overline{n_r(t)} = \frac{\lambda e^{-\lambda r}}{[1 + (\lambda\gamma_0/k)]^{2k+2}} \left[t - r - 2\gamma_0 \left(\frac{k+1}{k + \lambda\gamma_0} \right) \right], \quad \dots\dots\dots (3.9)$$

$$\overline{x_r(t)} = \frac{e^{-\lambda r}}{[1 + (\lambda\gamma_0/k)]^{2k+2}} \left[t - r - 2\gamma_0 \left(\frac{k+1}{k + \lambda\gamma_0} \right) - \lambda^{-1} \right], \quad \dots\dots\dots (3.10)$$

$$\mu_n^2 = 1 - \frac{2\lambda e^{-\lambda r}}{[1 + (\lambda\gamma_0/k)]^{2k+2}} \left[r + 2\gamma_0 \left(\frac{k+1}{k + \lambda\gamma_0} \right) \right] + \text{order } [(2\gamma_0 + r)/t], \quad \dots (3.11)$$

$$\mu_x^2 = (2/\lambda) \left\{ 1 - \frac{e^{-\lambda r}}{[1 + (\lambda\gamma_0/k)]^{2k+2}} \left[1 + \lambda r + 2\lambda\gamma_0 \left(\frac{k+1}{k + \lambda\gamma_0} \right) \right] + \text{order } \left(\frac{2\gamma_0 + r}{t} \right) \right\}. \quad \dots\dots\dots (3.12)$$

It can be seen by inspection that these expressions reduce to the ones of the O'Ceallaigh model in the limit $k \rightarrow \infty$. Furthermore, equation (3.6) is still valid and can again be used for a quick determination of the parameter λ .

* Strictly speaking, the mean length of individual gaps is the average value of (x/n) , not the ratio of the averages x/n . However, if the fluctuations are small the difference is insignificant. This is always so for long enough cell lengths t .

These expressions differ from the ones obtained by Happ, Hull, and Morrish because we concentrate our attention on different track parameters. In particular, the total excess gap length $x_r(t)$ was not considered by them (Happ, Hull, and Morrish 1952). We shall show in Section IV of this paper that it is impossible to fit both the total gap length and the mean number of gaps with values of γ_0 and k in agreement with experimental measurements on the emulsion. Hence the model of Happ, Hull, and Morrish will turn out to be inadequate. This fact could not have been seen from a study of the fluctuations in $n(t)$; the fluctuation parameter μ_n agrees with experiment.

(c) *A Modified O'Ceallaigh Model*

As a first step towards taking into account the "graininess" of the emulsion before development, we can use the survival probability (1.3) rather than (1.1). The modified survival probability takes cognizance of the fact that each crystal has a non-zero radius, and hence two crystals can never be closer together than some minimum distance α ; α corresponds to the diameter (not the radius) of the crystals before development. Since this model will also fail to fit the data, we shall give only some of the results, namely the coefficients of t in $n_r(t)$ and $x_r(t)$; these are:

$$n_r(t) = \frac{\lambda t}{1 + \lambda \alpha} \exp [\lambda(\alpha - 2\gamma - r)] + \text{constant}, \quad \dots \dots (3.13)$$

$$x_r(t) = \frac{t}{1 + \lambda \alpha} \exp [\lambda(\alpha - 2\gamma - r)] + \text{constant}. \quad \dots \dots (3.14)$$

The constant terms in each case are of relative order of magnitude $(2\gamma + r)/t$. We observe that equation (3.6) is still valid.

(d) *A Modification of the Model of Happ, Hull, and Morrish*

This model has the same survival probability (1.3), as the preceding model, but assumes a distribution in final grain radius $G(\gamma)$. It turns out to be convenient to use the following modification of the distribution (3.7):

$$\left. \begin{aligned} G(\gamma) &= 0, & \text{for } \gamma < \frac{1}{2}\alpha, \\ &= (k!)^{-1} (k/\gamma_0)^{k+1} (\gamma - \frac{1}{2}\alpha)^k \exp [-k(\gamma - \frac{1}{2}\alpha)/\gamma_0], & \text{for } \gamma > \frac{1}{2}\alpha. \end{aligned} \right\} \dots (3.15)$$

Physically this means that we assume that grains never shrink during the process of development. Since the mean grain radius after development, γ , exceeds the initial crystal radius, $\frac{1}{2}\alpha$, by a considerable margin, and since the distribution function $G(\gamma)$ is quite sharp in any case, the few grains which perhaps shrink during development have no appreciable influence on the final results. The mathematical evaluation of the various integrals is simplified greatly by (3.15). The relationship of the parameters γ_0 and k to the measured distribution of sizes of single grains is

$$\gamma_0 = \gamma - \frac{1}{2}\alpha - \frac{\overline{\gamma^2} - (\overline{\gamma})^2}{\overline{\gamma} - \frac{1}{2}\alpha}, \quad k = \frac{\gamma_0}{\gamma - \frac{1}{2}\alpha - \gamma_0}. \quad \dots \dots (3.16)$$

For the same reasons as before, we give only the coefficients of t in the mean values $\overline{n_r(t)}$ and $\overline{x_r(t)}$; these are:

$$\overline{n_r(t)} = \frac{\lambda t e^{-\lambda r}}{(1 + \lambda \alpha)[1 + (\lambda \gamma_0/k)]^{2k+2}} + \text{constant}, \quad \dots \quad (3.17)$$

$$\overline{x_r(t)} = \frac{t e^{-\lambda r}}{(1 + \lambda \alpha)[1 + (\lambda \gamma_0/k)]^{2k+2}} + \text{constant}, \quad \dots \quad (3.18)$$

Again the relative contribution of the constant terms is of order $(2\gamma_0 + r)/t$; again equation (3.6) can be used to find λ . Although (3.17) and (3.18) look very similar to (3.9) and (3.10), it must be remembered that the parameters γ_0 and k determined from (3.16) are quite different from γ_0 and k determined from (3.8), for the same emulsion.

(e) The Model of Herz

This model goes to the opposite extreme in treating the "graininess" of the emulsion. It is characterized by the survival probability (1.2) together with the assumption of a unique final grain radius γ^* . We introduce the integer R as follows:

$$R = \text{largest integer less than } (2\gamma + r)/\beta. \quad \dots \quad (3.19)$$

We then get:

$$\overline{n_r(t)} = p(1-p)^R (t/\beta - R), \quad \dots \quad (3.20)$$

$$\overline{x_r(t)} = p(1-p)^R [\bar{w}(t/\beta - R) - \beta(1-p)/p^2], \quad \dots \quad (3.21)$$

where the "mean gap width" \bar{w} is defined by

$$\bar{w} = \beta R - (2\gamma + r) + (\beta/p). \quad \dots \quad (3.22)$$

The fluctuations in this model are given by

$$\mu_n^2 = 1 - (2R+1)p(1-p)^R + \text{order } (2\gamma + r/t), \quad \dots \quad (3.23)$$

$$\begin{aligned} \mu_x^2 = \frac{w_0^2}{w} [1 - pq^R(1+2R)] + 2\beta(w_0/\bar{w})q^{R+1}[1+2R-4Rq-(q^2/p)] \\ + (\beta^2/\bar{w})(q/p)[1+2(q/p)-q^{R+1}(3+2R+2q/p)], \quad \dots \quad (3.24) \end{aligned}$$

where we have used the abbreviations

$$q = 1 - p, \quad \dots \quad (3.25)$$

and

$$w_0 = \lim_{p \rightarrow 1} \bar{w} = (R+1)\beta - (2\gamma + r). \quad \dots \quad (3.26)$$

In the model of Herz the mean gap width \bar{w} does *not* approach zero as the activation probability p comes close to unity; rather it approaches the length

* The quantity " γ " of Herz and Davis (1955) equals our $2\gamma/\beta$. For the special case $r=0$, that is, all gaps counted, our R equals the Γ of Herz.

of the "zero order gap", w_0 , expression (3.26). This is a characteristic difference between the Herz model and all models discussed so far. The difference can be traced directly to the assumption of Herz that the undeveloped crystals are spaced in a uniform lattice, that is, to the special form (1.2) of the survival probability $Q(t)$. This means that the grains either merge into a blob, or else have gap spacings equal to

$$w = w_0 + m\beta, \quad m = \text{integral}. \quad \dots\dots\dots (3.27)$$

As the track becomes more dense, the gaps move farther and farther apart, and a larger proportion of them become "zero order" gaps, that is, gaps with width $w = w_0$, $m = 0$ in (3.27). In the limit of a very dense track, all gaps are zero order gaps.

In the continuum models discussed so far, the mean gap width is given by equation (3.6), and thus approaches zero for very dense tracks (λ approaching infinity). The difference between the continuum models and the Herz model is unimportant for thin tracks, but becomes increasingly more important as the track gets heavier.

(f) *A Modification of the Herz Model*

This modification consists in retaining the survival probability (1.2) of the Herz model, but allowing the final grain size to vary according to a statistical distribution function $G(\gamma)$. The function $G(\gamma)$ is chosen to be (3.15), with α replaced by β , again because the integrals can be done in closed form, and the precise form of $G(\gamma)$ is unimportant in any case. The integrations with this model are not altogether trivial, and are discussed in some detail in Appendix I. Since the results are rather complicated, we shall restrict ourselves here to giving the coefficients of t in the mean values for $r=0$, that is, in \bar{n}_0 and \bar{x}_0 . If the model checks sufficiently well against experiment (which has not been determined as yet) the fluctuations can also be computed by the methods given in Appendix I. The mean values, omitting constant terms, are

$$\bar{n}_0(t) = p \left[1 - p \sum_{m=0}^{2k+1} \frac{(k\beta/\gamma_0)^m}{m!} T_m(y) \right] \left(\frac{t}{\beta} \right) + \text{constant}, \quad \dots\dots (3.28)$$

where k and γ_0 are the parameters in $G(\gamma)$, equation (3.15), β is the lattice constant in the Herz model, $T_m(y)$ is a function defined in Appendix I, and y is an abbreviation for

$$y = (1-p) \exp(-k\beta/\gamma_0). \quad \dots\dots\dots (3.29)$$

The mean total gap length, counting all gaps, is

$$\begin{aligned} \bar{x}_0(t) = & \left[1 - p - \frac{2(k+1)\gamma_0 p}{k\beta} + \frac{\gamma_0 p^2}{k\beta} \sum_{m=0}^{2k+1} (2k+2-m) \right. \\ & \left. \times \frac{(k\beta/\gamma_0)^m}{m!} T_m(y) \right] t + \text{constant}. \quad \dots\dots\dots (3.30) \end{aligned}$$

Numerical evaluation of these expressions and comparison against experimental data will form the subject of a later publication.

IV. COMPARISON WITH EXPERIMENT

The purpose of this paper has been to derive and state the theoretical formulas necessary to analyse experimental data. A detailed analysis of data on tracks in nuclear research emulsions will be presented in a later publication. However, it is perhaps of some interest to point out one qualitative result of this analysis at this stage. This is the inadequacy of the continuum models of track formation, that is, of all models which replace the survival probability by a completely smoothed-out function (usually an exponential).

In order to obtain a quick test of the various models, we choose one rather dense track in a G5 emulsion.* For this track the experimental values of the mean number of gaps and the total gap length are ($r=0$, i.e. all gaps are counted)

$$\left. \begin{aligned} \bar{n} &= 0.383t \quad (t \text{ in microns}), \\ \bar{x} &= 0.177t. \end{aligned} \right\} \dots\dots\dots (4.1)$$

In the same emulsion, the mean radius of developed single grains was 0.325μ . We first compare against the continuum models ((a), (b), (c), and (d) of Section III). In all these models the parameter λ is given by equation (3.6), so that

$$\lambda = 2.16_5 \mu^{-1}. \dots\dots\dots (4.2)$$

We now insert this value of λ and the grain size $\gamma = 0.325 \mu$ into formula (3.1a) to get, ignoring the constant term,

$$\bar{n} = 0.530t, \quad \text{O'Ceallaigh model.} \dots\dots\dots (4.3)$$

Since this is quite different from (4.1), we turn to the model of Happ, Hull, and Morrish. For this emulsion, reasonable values of γ_0 and k are: $\gamma_0 = 0.300 \mu$, $k = 10$. Substitution in (3.9) gives, again ignoring the constant term,

$$\bar{n} = 0.543t, \quad \text{model of Happ, Hull, and Morrish} \dots\dots (4.4)$$

Hence the statistics of final grain sizes has made little difference (as is to be expected for $k=10$) and in so far as it had an effect, the effect went in the wrong direction. We therefore turn to the modifications which take into account the initial crystal sizes to a first approximation (model C of Section I). In this emulsion, a reasonable value of β is 0.290μ . Substitution in (3.13) gives

$$\bar{n} = 0.616t, \quad \text{modified O'Ceallaigh model.} \dots\dots\dots (4.5)$$

Thus the modified O'Ceallaigh model is even worse than the original one. Finally, we put in the statistics for the final grain radii, this time with $\gamma_0 = 0.125 \mu$ and $k = 2.8$. The result is the worst yet:

$$\bar{n} = 0.652t, \quad \text{modified model of Happ, Hull, and Morrish.} \dots\dots (4.6)$$

* The data for this track were supplied to the author by Dr. A. J. Herz.

While we have picked only one track for this comparison, it should be noted that the results are typical for all thick tracks in nuclear research emulsions. Of course, the theory could be brought in agreement with experiment by choosing a different value of γ . To see what γ would have to be, we take the most favourable case (i.e. the one which came nearest to agreement before), that of the O'Ceallaigh model. We use λ as in (4.2) and determine γ from (4.1) and (3.1a). The result is

$$\gamma = 0.40 \mu. \quad \dots\dots\dots (4.7)$$

The experimental distribution of radii of single grains depends on the resolving power of the microscope, the adjustment of the microscope, the state of the emulsion, and the particular observer. For any combination of optical conditions and observer, however, the average grain diameter is a clearly defined quantity, and two adjacent grains will be counted as a single blob if their boundaries, as seen by that particular observer under these particular conditions, overlap. Hence, if the blob counting and gap measuring is carried out under the same conditions and by the same observer as the measurement of developed grain sizes, then the theory should apply, and the optical conditions and peculiarities of the observer need not be considered.

The same observer who measured the track listed in (4.1) also measured, in the same plate and under the same optical conditions, the distribution of blob sizes in a thin track, where most blobs correspond to single grains. The distribution of blob sizes in this thin track has a clearly defined peak corresponding to single-grain blobs. The most probable grain diameter is 0.65μ , and 43 per cent. of the grains are contained in the peak between 0.62 and 0.68μ . A grain diameter of 0.80μ is definitely in the tail of the distribution; the distribution function $G(\gamma)$ has dropped to about 1/7th of its peak value by the time grain diameters $2\gamma = 0.80 \mu$ are reached.

Hence we conclude that the models which ignore the approximate lattice structure of the spatial distribution of undeveloped crystals in the emulsion are definitely in disagreement with experimental data. For the same track already discussed, Herz had no trouble in fitting the data to his model. Furthermore, Herz was able to fit several tracks in the same emulsion using identical values of lattice spacing β and final grain radius γ , varying only the activation probability p .

A more detailed comparison against experimental data will form the subject of a later publication.

V. ACKNOWLEDGMENT

The author would like to thank Dr. A. J. Herz for many valuable discussions.

VI. REFERENCES

- BELLMAN, R., and HARRIS, T. E. (1948).—*Proc. Nat. Acad. Sci., Wash.* **34**: 601.
 HAPP, W. W., HULL, T. E., and MORRISH, A. H. (1952).—*Canad. J. Phys.* **30**: 699.
 HERZ, A. J., and DAVIS, G. (1955).—*Aust. J. Phys.* **8**: 129.
 JÁNOSSY, L. (1950).—*Proc. Phys. Soc. Lond.* **A 63**: 241.
 O'CEALLAIGH, C. (1953).—Rep. Cosmic Ray Congress, Bagnères-de-Bigorre, p. 73.

APPENDIX I

Some Methods for Evaluating the Integrals

The integrals (2.34)–(2.40) must be evaluated, preferably in closed form, before the theory makes useful statements. In general there is no difficulty with (2.34) and (2.35), which are just the first two moments of the survival probability $Q(t)$. We shall therefore restrict ourselves to a discussion of the integrals g_{ij} , (2.36)–(2.40).

It turns out to be very convenient to introduce the Laplace transform of the survival probability :

$$P(s) = \int_0^\infty e^{-st} Q(t) dt, \quad \dots\dots\dots (\text{A1})$$

as well as its derivative

$$P'(s) = \int_0^\infty e^{-st} t Q(t) dt. \quad \dots\dots\dots (\text{A2})$$

Clearly the following relations hold

$$P(0) = Q_0, \quad P'(0) = -Q_1. \quad \dots\dots\dots (\text{A3})$$

Consider now the integral g_{00} , expression (2.36). We observe that the integration over t simply yields the function $Q(\gamma_1 + \gamma_2 + r)$, for which we write

$$Q(\gamma_1 + \gamma_2 + r) = (2\pi i)^{-1} \int_C ds \exp [s(\gamma_1 + \gamma_2 + r)] P(s). \quad \dots\dots (\text{A4})$$

The contour in (A4) is the standard Bromwich contour, to the right of all poles of $P(s)$. We then interchange orders of integration over γ_1 , γ_2 , and s . We introduce the Laplace transform of the distribution-in-final-grain-radius $G(\gamma)$:

$$H(s) = \int_0^\infty e^{-s\gamma} G(\gamma) d\gamma, \quad \dots\dots\dots (\text{A5})$$

in order to get

$$g_{00}(r) = (2\pi i)^{-1} \int_C ds e^{sr} H^2(-s) P(s), \quad \dots\dots\dots (\text{A6})$$

where the contour C must be restricted in such a way that $H(-s)$ represents a convergent integral rather than the analytic continuation of a convergent integral. That is, *the contour C must lie to the left of all poles of $H(-s)$ as well as to the right of all poles of $P(s)$* . These conditions may turn out to be contradictory, in which case this method of evaluating the integrals fails : but so far it has been quite easy to satisfy both conditions simultaneously.

Next consider g_{01} , equation (2.37). We introduce for the moment the notation

$$T = \gamma_1 + \gamma_2 + r, \quad \dots\dots\dots (\text{A7})$$

and we ask for the Laplace transform of the function of T obtained by performing the integration over t , that is,

$$U(s) = \int_0^\infty dT \int_T^\infty dt e^{-st} t q(t). \quad \dots\dots\dots (\text{A8})$$

We interchange orders of integration to get

$$U(s) = s^{-1} \int_0^{\infty} dt (1 - e^{-st}) tq(t),$$

and integrate by parts to obtain

$$U(s) = s^{-1} [Q_0 - P(s) - sP'(s)]. \quad \dots\dots\dots (\text{A9})$$

From here on the steps are just as before, leading to the result

$$g_{01}(r) = (2\pi i)^{-1} \int_C ds e^{sr} H^2(-s) \frac{Q_0 - P(s) - sP'(s)}{s}. \quad \dots\dots (\text{A10})$$

Contrary to the first impression, the integrand of (A10) is regular at $s=0$; this can be seen by using (A3), or more directly by observing that (A8) is by definition a convergent integral when $s=0$.

The same approach can be used on the other three integrals g_{ij} and yields :

$$g_{10}(r) = (2\pi i)^{-1} \int_C ds e^{sr} H^2(-s) \frac{Q_0 - P(s)}{s}, \quad \dots\dots\dots (\text{A11})$$

$$g_{11}(r) = (2\pi i)^{-1} \int_C ds e^{sr} H^2(-s) \frac{P(s) + sP'(s) - Q_0 + 2sQ_1}{s^2}, \quad \dots\dots (\text{A12})$$

$$g_{20}(r) = (2\pi i)^{-1} \int_C ds e^{sr} H^2(-s) \frac{2[P(s) - Q_0 + sQ_1]}{s^2}. \quad \dots\dots\dots (\text{A13})$$

In all cases the integrand is regular at $s=0$. In most cases the contour can be deformed towards the left, so that the relevant poles are the poles of $P(s)$ (which are of course at the same points as the poles of $P'(s)$). In the continuum models, with survival probabilities (1.1) or (1.3), $P(s)$ is very simple and has only one pole, at $s=-\lambda$.

The rest of this appendix is devoted to the evaluation of the integrals in the model of Herz and in the generalization of that model discussed in Section III.

We introduce the notation

$$q = 1 - p, \quad \dots\dots\dots (\text{A14})$$

where p is the probability of activating a traversed crystal. The Laplace transform of the survival probability $Q(t)$, expression (1.2), is

$$\begin{aligned} P(s) &= \sum_{k=0}^{\infty} q^k \int_{k\beta}^{(k+1)\beta} e^{-st} dt \\ &= s^{-1} [1 - e^{-s\beta}] \sum_{k=0}^{\infty} [qe^{-s\beta}]^k. \quad \dots\dots\dots (\text{A15}) \end{aligned}$$

We can sum the infinite series to get the form

$$P(s) = \frac{1 - e^{-s\beta}}{s[1 - qe^{-s\beta}]}. \quad \dots\dots\dots (\text{A16})$$

However, we shall find that (A15) is often a more convenient form to work with. The poles of $P(s)$ are located at

$$s_n = -\beta^{-1}[\ln(1/q) + n(2\pi i)], \quad n = \text{integer}. \quad \dots\dots (\text{A17})$$

Since q is by definition less than unity, it suffices to take our contour C to the right of the origin.

To illustrate the method, let us first evaluate $g_{00}(r)$ for the Herz model. In this model $G(\gamma)$ is a delta function, so that

$$H(s) = e^{-s\gamma_0}. \quad \dots\dots\dots (\text{A18})$$

Substitution of (A15) and (A18) into (A6) yields

$$\begin{aligned} g_{00}(r) &= (2\pi i)^{-1} \int_C ds \exp[s(r+2\gamma_0)] \frac{1 - \exp(-s\beta)}{s} \sum_{k=0}^{\infty} q^k \exp(-ks\beta) \\ &= (2\pi i)^{-1} \int_C ds s^{-1} \exp[s(r+2\gamma_0)] \left\{ 1 - p \sum_{k=1}^{\infty} q^{k-1} \exp(-ks\beta) \right\}. \\ &\quad \dots\dots\dots (\text{A19}) \end{aligned}$$

We now interchange the order of integration and summation; this is possible since the infinite series converges absolutely and uniformly along the entire contour C . The contributions of the various terms k depend on the magnitude of k . First consider small values of k , for which the coefficient in the exponential $\exp[s(r+2\gamma_0-k\beta)]$ is positive. We then close the contour to the left and obtain the contribution from the pole at $s=0$ only. This contribution is $-p(q)^{k-1}$. But when k gets large enough that the coefficient in the exponential is negative, we close the contour to the right, and get zero. The critical value of k is precisely the integer R defined by (3.19). Hence we get

$$g_{00}(r) = 1 - p \sum_{k=1}^R q^{k-1} = 1 - p \frac{1 - q^R}{1 - q} = 1 - (1 - q^R) = q^R. \quad \dots\dots (\text{A20})$$

Of course, this integral could have been evaluated much more directly from first principles, but this derivation was intended to show an application of the method of this appendix to a particularly simple case.

We now turn to the modified model of Herz, where the integrals are somewhat less trivial. The distribution-in-final-grain-radius $G(\gamma)$ is given by (3.15), and its Laplace transform is

$$H(s) = \frac{\exp(-\frac{1}{2}s\beta)}{[1 + (\gamma_0 s/k)]^{k+1}}. \quad \dots\dots\dots (\text{A21})$$

For large values of k this reduces to

$$\lim_{k \rightarrow \infty} H(s) = \exp[-s(\gamma_0 + \frac{1}{2}\beta)]. \quad \dots\dots\dots (\text{A22})$$

This is of course the same as (A18) if we remember the altered definition of the parameter γ_0 .

We evaluate $g_{00}(r)$ explicitly for the special case $r=0$; it should be noted that it would be very difficult to get a closed result without the use of these Laplace transform methods. Substitution of (A15) and (A21) into (A6) gives

$$\begin{aligned} g_{00}(r) &= (2\pi i)^{-1} \int_C ds \frac{[1 - \exp(-s\beta)] \exp[s(r+\beta)]}{s[1 - (\gamma_0 s/k)]^{2k+2}} \sum_{n=0}^{\infty} q^n \exp(-ns\beta) \\ &= (2\pi i)^{-1} \int_C ds \frac{\exp(sr)}{s[1 - (\gamma_0 s/k)]^{2k+2}} \left\{ \exp(s\beta) - p \sum_{n=0}^{\infty} q^n \exp(-ns\beta) \right\}. \end{aligned} \quad \text{..... (A23)}$$

In order to get formulas of reasonable length, we restrict ourselves now to the special case of the limit $r \rightarrow 0$. It should be emphasized, however, that non-zero values of r can be handled by the same methods; the results are merely a little more involved. If we assume that r is an infinitesimal positive number, we see that we must close the contour to the left for the $e^{s\beta}$ term and for the term $n=0$ in (A23), to the right for all terms $n \geq 1$. The contour C runs to the left of the origin but to the right of the pole

$$s = \sigma \equiv k/\gamma_0. \quad \text{..... (A24)}$$

The contribution of the pole at $s=0$ is simple; let us now consider a typical term $n \geq 1$ and its contribution, which arises from the pole $s=\sigma$. This contribution is the negative of the residue at $s=\sigma$, that is, setting $r=0$,

$$\begin{aligned} &+ p q^n (-\sigma)^{2k+2} \frac{1}{(2k+1)!} \left\{ \frac{d^{2k+1}}{ds^{2k+1}} [s^{-1} e^{-ns\beta}] \right\}_{s=\sigma} \\ &= -p \sum_{m=0}^{2k+1} \frac{(n\beta\sigma)^m}{m!} [q e^{-\beta\sigma}]^n. \end{aligned}$$

We therefore get

$$g_{00}(0) = q - p \sum_{n=1}^{\infty} \sum_{m=0}^{2k+1} \frac{(n\beta\sigma)^m}{m!} [q e^{-\beta\sigma}]^n. \quad \text{..... (A25)}$$

At this stage we interchange the order of the two summations, and evaluate the infinite sum over n explicitly. We define

$$y = q \exp(-\beta\sigma) = (1-p) \exp(-k\beta/\gamma_0), \quad \text{..... (A26)}$$

and

$$T_m(y) = \sum_{n=0}^{\infty} n^m y^n = (y d/dy)^m \left(\frac{1}{1-y} \right), \quad \text{..... (A27)}$$

in order to get

$$g_{00}(0) = q - p \sum_{m=0}^{2k+1} \frac{(\beta\sigma)^m}{m!} [T_m(y) - \delta_{m,0}],$$

where $\delta_{m,0}$ is the Kronecker delta; the final result is

$$g_{00}(0) = 1 - p \sum_{m=0}^{2k+1} \frac{(k\beta/\gamma_0)^m}{m!} T_m(y). \quad \text{..... (A28)}$$

This result is not overly simple, but it would of course be unreasonable to expect a simple answer, since we must get a continuous and differentiable function of β/γ_0 which in the limit $k \rightarrow \infty$ approaches the discontinuous function (A20).

It may be worth while to say a few words about the functions $T_m(y)$ defined by (A27). Their calculation can be simplified by expressing the operators $(y d/dy)^m$ as linear combinations of the operators

$$D_r = y^r (d/dy)^r. \quad \text{..... (A29)}$$

We write

$$(y d/dy)^m = \sum_{r=1}^m a_{mr} D_r, \quad m \neq 0, \quad \text{..... (A30)}$$

and obtain the recursion relations

$$\begin{aligned} a_{m+1, r} &= a_{m, r-1} + a_{m, r}, & r \neq 1, r \neq m+1, \\ a_{m+1, 1} &= a_{m+1, m+1} = 1. & \text{..... (A31)} \end{aligned}$$

The first few coefficients a_{mr} are given in Table 1. We now introduce (A30) into (A27) to get

$$\begin{aligned} T_m(y) &= (1-y)^{-1} \sum_{r=1}^m r! a_{mr} \left(\frac{y}{1-y} \right)^r, & m \neq 0, \\ T_0(y) &= (1-y)^{-1}. & \text{..... (A32)} \end{aligned}$$

TABLE 1
VALUES OF THE COEFFICIENTS a_{mr} IN EQUATION (A32)

$m \backslash r$	1	2	3	4	5	6	7
1	1						
2	1	1					
3	1	3	1				
4	1	7	6	1			
5	1	15	25	10	1		
6	1	31	90	65	15	1	
7	1	63	301	350	140	21	1

Expression (A32) together with Table 1 gives a practical way of calculating $T_m(y)$. It should be noted that all the terms in the sum over r are positive, so that there is no danger of near cancellation of two large numbers of opposite sign. There is, however, appreciable cancellation of this type between the "1" in (A28) and the sum over m , so that each $T_m(y)$ must be calculated to considerably higher accuracy than the accuracy of the final result.

An entirely similar derivation gives

$$g_{10} = \beta q/p - (2k+2)\gamma_0/k + (p\gamma_0/k) \sum_{m=0}^{2k+1} (2k+2-m) \frac{(k\beta/\gamma_0)^m}{m!} T_m(y). \quad \text{.. (A33)}$$

These results are sufficient to determine the coefficients of t in the mean number of gaps and the mean total gap length, quoted in Section III.

THE INFLUENCE OF NOISE ON RADAR METEOR OBSERVATIONS

By C. D. ELLYETT* and G. J. FRASER*

[*Manuscript received February 1, 1955*]

Summary

An experimental investigation has been made of the minimum detectable echo power from meteor trails, using radar techniques at 69 Mc/s, with incoherent detection.

For optimum signal-to-noise ratio the pulse width must exceed the cathode-ray tube spot width. Extraterrestrial noise is predominant at 69 Mc/s, and the receiver noise figure is found to be of only second order importance. Variations in the effective aerial temperature introduce corresponding variations in the observed meteor rate, and must therefore be considered when meteor rates are being compared.

Using artificial echoes, a minimum detectable signal-to-noise ratio of 8 db is found. The most important parameter is the total received signal energy. Both these results are in agreement with published theories. No further reduction in the size of the smallest observable meteor is likely through receiver improvement. The important features of the receiver are described.

I. INTRODUCTION

In the radar detection of meteor trails the minimum detectable electron density, and therefore also the observed echo rate, depends on the parameters of the radar system (Kaiser 1953). Hence, if absolute echo rates and trail density comparisons between different observing stations are to be made, these parameters must be known accurately. The aerial gain and beam shape, the transmitter power, the wavelength, and the echo range can all be measured to a reasonable order of accuracy, but the minimum detectable echo power is not so readily obtained.

The accurate calculation of the minimum detectable signal power in a radar receiver (Norton and Omberg 1947; Goldman 1948; Lawson and Uhlenbeck 1950; Ross 1951; Spencer 1951) requires the knowledge of a large number of parameters which are not all easily determinable. Therefore, for the purpose of estimating meteor magnitudes, it was preferable to find the minimum detectable signal power by direct measurement.

The radar system for the detection of meteors uses conventional incoherent pulse techniques, and includes a 69 Mc/s transmitter with a peak power of 80 kW and a pulse recurrence frequency of 150 sec^{-1} , an aerial array with a gain of 120 relative to an isotropic source and half-power beam widths of 14° and 22° , and a receiver with a noise figure of 2.0. The signals are displayed on an intensity-modulated cathode-ray tube, and recorded on film moving continuously at right angles to the time base at a speed of $1.23 \times 10^{-2} \text{ cm sec}^{-1}$. Further details of the receiver are given in Appendix I.

* Canterbury University College, Christchurch, N.Z.

II. BANDWIDTH LIMITATIONS

The first bandwidth limitation is well known. With incoherent detection, the lowest possible value of the minimum detectable signal requires optimum setting of the pre-detection bandwidth (Goldman 1948; Spencer 1951). Video bandwidths wider than the optimum have negligible effect, but a bandwidth less than the pre-detection width degrades the signal-to-noise ratio.

In addition, however, because of the finite spot size the cathode-ray tube behaves as a low-pass filter (Ridenour 1947), since events occurring within the time taken for the spot to travel a distance equal to its own diameter are not resolved. It is therefore desirable to have the cathode-ray tube bandwidth wider than the pre-detection bandwidth, i.e. the cathode-ray tube spot width should be shorter than the signal pulse.

Meteor trails occur at a height of approximately 100 km. Hence with a low-elevation radiated beam a maximum range of about 900 km is required for the complete recording of echoes. The spot width thus represents more

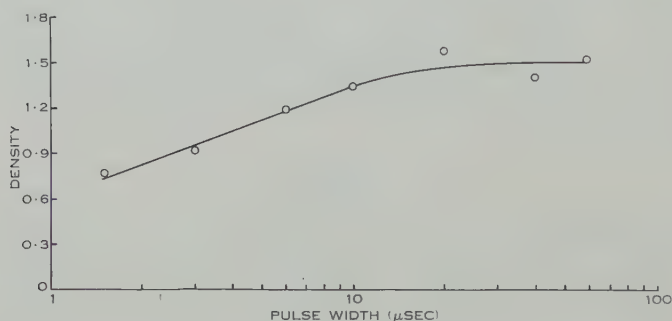


Fig. 1.—Variation of trace density with pulse width. (Pulse voltage = 15 V.)

kilometres range, or microseconds, than in conventional radar systems, and *pulse widths longer than usual must be used* if the optimum signal-to-noise ratio is to be achieved.

A series of measurements was made by applying to the cathode-ray tube grid a pulse of adjustable width, but of insufficient voltage to saturate the screen. The resultant film densities were measured on a microdensitometer (Tait and Chalklin 1953). The results, shown in Figure 1, demonstrate that the low-pass filter action of the cathode-ray tube produces appreciable loss with pulses shorter than 20 μ sec width.

Microscope measurements were also made of short-duration meteor echoes, and gave a mean value of 56 μ sec for the spot width. (Statistically significant variations occurred between films because of the variation of spot size with the brilliance control setting.) This time interval is longer than the value given by the microdensitometer method. Soller, Starr, and Valley (1948), however, find that microscope readings of cathode-ray traces are about double those recorded by other methods which approach more closely the conditions of visual

observation. The present figures are in fair agreement with this conclusion, and a value of 20 μ sec was therefore taken as the minimum desirable pulse length.

III. EFFECTS OF EXTERNAL NOISE

The minimum detectable signal is reduced by decreasing the receiver noise level. Extraterrestrial noise, however, is always present, and the effective aerial temperature, for an aerial beam more than a few degrees wide, is always greater than 1500 °K at 69 Mc/s (Norton and Omberg 1947; Moxon 1949). A reduction in internal receiver noise, if the latter is already low, may therefore produce little improvement. The total effective noise input power to the receiver (P_n) is

$$P_n = (F_e + F - 1)kTB,$$

where F_e = the ratio of the effective aerial temperature (T_a) to the receiver temperature ($T = 290$ °K)

= external noise factor,

F = receiver noise figure,

k = Boltzmann's constant,

B = receiver bandwidth.

The meteor echo rate for sporadic meteors is inversely proportional to the minimum detectable signal voltage (Kaiser 1953), hence

$$(N_2/N_1)^2 = (F_e + F_1 - 1)/(F_e + F_2 - 1), \quad \dots \dots \dots (1)$$

where N_2 (or N_1) is the meteor echo rate for a receiver noise figure of F_2 (or F_1). It can be seen that, if F_e is large compared with F , variations in F should produce little difference in the echo rate.

This independence of N_2/N_1 and F was verified experimentally in the following manner. The sporadic meteor echo rate was measured for consecutive 19 min intervals, and echo rates of the order of 60 per hour were obtained. On alternate 19 min periods the receiver noise figures were 2.01 ± 0.02 and 6.3 ± 0.2 respectively, the higher noise figure being obtained by reducing the gain of the first stage of the receiver. The meteor echo rates obtained with these two noise figures were virtually unchanged, giving a ratio $N_2/N_1 = 0.9 \pm 0.2$, whereas in the absence of external noise ($F_e = 0$) the ratio would have been 0.44. It is therefore apparent that external noise has much more effect on the echo rate than receiver noise. Quite apart from any inherent variation in real meteor rate with time, due to the Earth's rotation and translation, the observed or apparent meteor rate must also be a function of the equivalent aerial temperature.

The ratio of the maximum to minimum aerial temperatures is of the order of 6 (Shain 1954) or 10 (Hey, Parsons, and Phillips 1948). The corresponding ratio of echo rates is 0.44 or 0.34, assuming a minimum temperature of 1500 °K and $F = 2.0$. It is therefore important that the aerial temperature be considered when comparing echo rates, even if these rates are obtained with the same equipment.

IV. MEASUREMENT OF THE MINIMUM DETECTABLE SIGNAL POWER

Because noise introduces random variations in the observation of a signal, various definitions of the minimum detectable signal ($P_{\min.}$) are possible. The value of $P_{\min.}$ adopted here will be that having a 90 per cent. probability of detection, as assumed by Lawson and Uhlenbeck (1950).

The minimum observable signal was determined experimentally for the receiver of known noise factor by the following method. A pulsed signal generator was used to produce artificial echoes of 1.0 sec duration, since this is the effective integrating time on the film. More than 200 echoes, random in amplitude, range, and echo separation time, were recorded photographically for each test. The fraction detected in each amplitude group was assumed to be

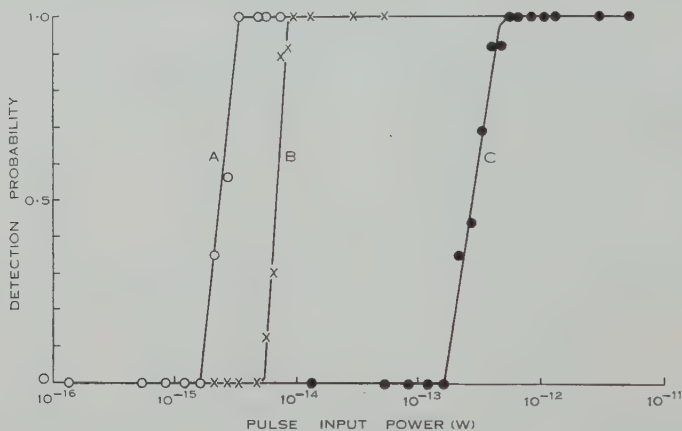


Fig. 2.—Detection probability as a function of pulse input power. The size of the points includes the estimated experimental error.

Curve A: I.F. bandwidth=63 kc/s; pulse width=26 μ sec;
 $P_{\min.}=0.30 \times 10^{-14}$ W (7.8 db above noise).

Curve B: I.F. bandwidth=300 kc/s; pulse width=8 μ sec;
 $P_{\min.}=0.85 \times 10^{-14}$ W (5.4 db above noise).

Curve C: I.F. bandwidth=300 kc/s; pulse width=3 μ sec;
 $P_{\min.}=40 \times 10^{-14}$ W (22.0 db above noise).

the detection probability of that amplitude. (Each echo was made to appear as a pair of dots or lines by displacing the time base slightly on alternate sweeps. The probability of random noise producing two pulses of large amplitude with the correct spacing is very small, so the number of chance guesses should be small, and has been neglected.)

The detection loss with pulse widths shorter than optimum (here about 20 μ sec) is clearly apparent from Figure 2. The minimum detectable power of 7.8 db above the noise level is in agreement with the experimental result of 8 db obtained by Gérardin (1954) for an intensity modulated radar system, and further confirms a theoretical analysis due to Ross (1951). When the receiver is connected to an aerial, $P_{\min.}$ will be 8 db above the total noise power

of $(F_e + F - 1)kTB$. An appreciable decrease in $P_{\min.}$ is only possible if coherent detectors are used (Goldman 1948; Smith 1951; Tucker and Griffiths 1953).

Curve A in Figure 2 was for optimum conditions. For curve B the pulse width was shorter, and the predetection bandwidth wider than the optimum suggested by Figure 1, but the pulse energies for these two curves in Figure 2 are equal within the experimental error. This confirms the results obtained by Goldman (1948), and by Woodward and Davis (1950), who analyse receiver performance in terms of information theory and find that the significant parameter is the ratio of the total received signal energy to the noise power per unit bandwidth. Their analyses assume coherent detection, but also hold for incoherent detection if the signal-to-noise ratio is large. The energy equivalence between curves A and B suggests that a signal power of 8 db above the noise is sufficiently large for the latter qualification to hold.

In the experiment described above the integrating time is of the order of 1.0 sec. but most meteor trail echoes do not last for this length of time at 69 Mc/s. Variation in the trail duration, if less than the integration time, will cause an inverse variation in $P_{\min.}$. Assuming an average echo duration of 0.1 sec, the value of $P_{\min.}$ in Figure 2 must be multiplied by 10.

V. THE MINIMUM DETECTABLE METEOR MAGNITUDE

Taking as average values for meteor trail detection an aerial temperature of 2000 °K, and a range of 480 km, together with the radar parameters described above, the minimum trail density ($\alpha_{\min.}$) is 2.4×10^{12} electrons per metre. From Kaiser's equation (1953) this value is equivalent to a zenithal magnitude of +9. Polarization effects are small and have been neglected in this calculation. A perfectly noiseless receiver would only reduce $\alpha_{\min.}$ to $2.1 \times 10^{12} \text{ m}^{-1}$. It therefore appears that $\alpha_{\min.}$ at 69 Mc/s can only be reduced further by increasing the radiated power density. This can be achieved either by increasing the mean transmitter power considerably, or by increasing the aerial gain. No further decrease in $\alpha_{\min.}$ is likely through receiver improvement.

Meteors show longer durations with decreasing radar frequency. This implies a greater integration time, and allows smaller meteors to be detected. Extraterrestrial noise, however, also increases with decreasing frequency, so the change in observed meteor rate on moving to lower frequencies can be expected to be less than that arising from the frequency effect alone.

In the comparison of trail densities and rates it would be preferable to specify the minimum detectable energy rather than the minimum detectable power.

VI. ACKNOWLEDGMENTS

This paper represents part of the New Zealand programme of meteor research financed jointly by the University of New Zealand, the Nuffield Foundation, and the New Zealand Department of Scientific and Industrial Research. The authors are indebted to Mr. L. Adams for his assistance with the experimental work.

VII. REFERENCES

- CRONEY, I. (1951).—*Proc. Inst. Radio Engrs., N.Y.* **39**: 807.
- GÉRARDIN, L. (1954).—*Onde Élect.* **34**: 67.
- GOLDMAN, S. (1948).—*Proc. Inst. Radio Engrs., N.Y.* **36**: 584.
- HEY, J. S., PARSONS, S. J., and PHILLIPS, J. W. (1948).—*Proc. Roy. Soc. A* **192**: 425.
- KAISER, T. R. (1953).—*Advanc. Phys.* **2**: 495.
- LAWSON, J. L., and UHLENBECK, G. E. (1950).—"Threshold Signals." (McGraw-Hill: New York.)
- MOXON, L. A. (1949).—"Recent Advances in Radio Receivers." (Cambridge Univ. Press.)
- NORTON, K. A., and OMBERG, A. C. (1947).—*Proc. Inst. Radio Engrs., N.Y.* **35**: 927.
- RIDENOUR, L. N. (1947).—"Radar System Engineering." (McGraw-Hill: New York.)
- ROSS, A. W. (1951).—*Wireless Engr.* **28**: 79.
- SHAIN, C. A. (1954).—*Aust. J. Phys.* **7**: 150.
- SMITH, R. A. (1951).—*Proc. Instn. Elect. Engrs.* IV **98**: 43.
- SOLLER, T., STARR, M. A., and VALLEY, G. E. (1948).—"Cathode-ray Tube Displays." (McGraw-Hill: New York.)
- SPENCER, R. E. (1951).—*J. Brit. Instn. Radio Engrs.* **11**: 435.
- TAIT, R. I., and CHALKLIN, F. C. (1953).—*J. Sci. Instrum.* **30**: 268.
- TUCKER, D. G., and GRIFFITHS, J. W. R. (1953).—*Wireless Engr.* **30**: 264.
- WALLMAN, H., MACNEE, A. B., and GADSDEN, C. P. (1948).—*Proc. Inst. Radio Engrs., N.Y.* **36**: 700.
- WOODWARD, P. M., and DAVIES, I. L. (1950).—*Phil. Mag.* **41**: 1001.

APPENDIX I

A 69 Mc/s receiver was constructed for the detection of echoes from meteor trails. The receiver noise has been reduced by using a cascode radio frequency amplifier (Wallman, MacNee, and Gadsden 1948). Provision is made for adjustment of the source impedance by using a Π -network to couple the aerial to the radio-frequency amplifier. The network capacitors are variable, and have been calibrated. Once the optimum source impedance and optimum detuning have been determined with the help of a noise generator, the network constants can be adjusted to suit a wide range of aerial impedances. Correct neutralization gave a noise figure of 1.6. Since such low values of receiver noise are unimportant, a noise figure of 2.0 has been used to allow for any small variations in input tuning and neutralization.

The relationship between signals differing largely in amplitude was required to be known for meteor velocity measurements. A logarithmic I.F. characteristic satisfies this requirement, and the circuit suggested by Croney (1951) was used as it is instantaneous and requires no critical matching of components. Five amplifying stages and six detectors are used, and the output is closely logarithmic over a range of at least 60 db. For linear receiver operation all detectors but the last are biased off by returning the cathodes to +25 V.

DIFFUSION COEFFICIENTS FROM THE RATE OF DECAY OF METEOR TRAILS

By A. A. WEISS*

[*Manuscript received December 29, 1954*]

Summary

The effective diffusion coefficient for a meteor trail is calculated from the theory of ambipolar diffusion and the physical constants of the upper atmosphere. The absolute value of the diffusion coefficient so calculated, and also its gradient with height, are confirmed by measurement of the rates of decay of a large number of meteor echoes of known heights. The individual values show considerable scatter, most of which is attributed to a regular diurnal variation in the value of the diffusion coefficient. Amplitude fluctuations in persistent echoes are also briefly discussed.

I. INTRODUCTION

The problem of the behaviour of ionized meteor trails after formation has been considered in detail by Greenhow (1950, 1952). In the second paper the distinction between echoes of short and long duration was drawn, and evidence was presented that the dissipation of the ionization is primarily due to diffusion processes. Fluctuations in the amplitudes of long-enduring echoes were ascribed to distortion of the initially uniform ionized column into two or more reflecting centres. The observations discussed by Greenhow were all obtained by pulse techniques on wavelengths of 4.2 and 8.4 m. The more recent observational data are contained in the comprehensive report by Kaiser (1953).

Huxley (1952) has shown that the durations of meteor echoes are of the order of magnitude to be expected from ambipolar diffusion of the trails and the known behaviour of ions and electrons in gases. In this paper a revised estimate of the diffusion coefficient in the meteor zone is obtained by an extension of Huxley's theory of ambipolar diffusion, and these theoretical values are compared with data obtained at Adelaide using the c.w. technique at a wavelength of 11.2 m, described by Robertson, Liddy, and Elford (1953).†

II. THE EFFECTIVE DIFFUSION COEFFICIENT

According to Kaiser (1953) the loss of electrons by recombination or by attachment is negligible except perhaps in the final decay of very persistent echoes, and the heat generated during the formation of the trail has little effect

* Department of Physics, University of Adelaide.

† *Note added in Proof.*—Theoretical calculations of the effective diffusion coefficient, essentially similar to those of Section II but without consideration of the height gradient, have been published by Massey and Sida (1955) after preparation of this paper. Some measurements by Greenhow and Neufeld (1955) on a much smaller sample of echoes have also appeared.

upon its subsequent behaviour. It will therefore be assumed that the dissipation of meteor ionization is due to ambipolar diffusion alone.

The radial distribution of electrons in an initially linear concentration, found by solution of the diffusion equation, is

$$n_c = (\alpha/4\pi Dt) \exp(-r^2/4Dt),$$

where D is the effective diffusion coefficient and α the line density of electrons (Huxley 1952). The distribution is therefore Gaussian. If the line density α is sufficiently small, the incident wave penetrates throughout the column and the echo amplitude decays exponentially from an initial value A_0 according to the law

$$A = A_0 \exp(-16\pi^2 Dt/\lambda^2). \quad \dots\dots\dots (1)$$

For $\alpha < 2.4 \times 10^{12}/\text{cm}$, A_0 is the maximum amplitude according to the Lovell-Clegg scattering formula, although, if predicted resonance effects (Herlofson 1951; Kaiser and Closs 1952) are present, this is true only for parallel polarization. For transverse polarization the amplitude near the beginning of the echo is enhanced by the resonance and in this event the exponential decay law (1) will cease to apply. When $\alpha > 2.4 \times 10^{12}/\text{cm}$ (persistent echoes) the law (1) applies only to the final stage of decay of the echo, when the electron density has become sufficiently dilute.

The effective diffusion coefficient for electrons in a meteor trail may be written

$$D \sim D_i(1 + T_e/T_i),$$

where D_i is the diffusion coefficient for positive ions and T_e and T_i are the electron and ion temperatures respectively. Since $T_e \sim T_i$ (Huxley 1952), then $D \sim 2D_i$. The coefficient D_i may be evaluated, without recourse to the theoretical formulae of kinetic theory, following an unpublished method due to Huxley and Robertson.

The diffusion coefficient D_i and the mobility a_i of a group of singly charged ions are connected by the relation

$$D_i = a_i kT/e, \quad \dots\dots\dots (2)$$

where k is Boltzmann's constant and e the electronic charge. The mobilities in nitrogen of the positive ions of a large number of metals have been measured (Tyndall 1938) and it is found that at a pressure of 760 mm Hg and a temperature of 18 °C the mobilities of these ions in most instances have values lying between 2 and 3 $\text{cm}^2 \text{sec}^{-1} \text{V}^{-1}$, although the masses of the ions differ greatly. Further, Pearce (1936) found that, for nitrogen and caesium ions moving in nitrogen at constant temperature, a change in temperature of 400 °C changed a_i by a factor of 2 only. If this behaviour is accepted as typical, then for temperatures between 200 and 250 °K and at a pressure of 760 mm Hg, the values of a_i for meteor ions are about $2.5 \text{ cm}^2 \text{sec}^{-1} \text{V}^{-1} = 7.5 \times 10^2 \text{ cm}^2 \text{sec}^{-1} (\text{e.s.u.})^{-1}$. Since a_i is inversely proportional to the molecular concentration, the value of D is

$$\begin{aligned} D &= 2D_i = 1.50 \times 10^3 (760/p) (T^2/291) k/e \\ &= 1.13 \times 10^{-3} T^2/p \quad \text{cm}^2/\text{sec}. \quad \dots\dots\dots (3) \end{aligned}$$

Using values of T and p found by means of rockets (Rocket Panel 1952) the variation of D with height may be found from (3). This relation, sketched in Figure 3, is to be compared with measured values of D .

The effect of the Earth's magnetic field, not so far considered, is to modify the diffusion coefficient for electrons, D_e , without affecting the more massive positive ions. For electrons moving parallel to the magnetic field in the absence of positive ions, the diffusion coefficient is $D_{||} = kT/mv \gg D_i$. Here v is the collisional frequency and m the electronic mass. The rate of diffusion of the meteor trail in this direction is clearly controlled by the positive ions at all heights.

On the other hand, for electrons alone moving transverse to the magnetic field,

$$D_{\perp} = D_{||} v^2 / (v^2 + \omega^2) \\ \sim D_{||} v^2 / \omega^2, \quad \text{if } \omega^2 / v^2 \gg 1.$$

With the gyro-frequency $\omega \sim 10^7$ radians/sec, and using the expression for the collisional frequency $v = 9.36 \times 10^7 p$ (Crompton, Huxley, and Sutton 1953),

$$D_{\perp} = 1.44 \times 10^5 T p, \quad p \text{ in mm Hg.} \quad \dots \dots \dots (4)$$

Comparing (4) with (3),

$$D_{\perp} / D = 1.28 \times 10^8 p^2 / T. \quad \dots \dots \dots (5)$$

Again using the rocket data, it is found that $D_{\perp} = D_i = D$ at a height of approximately 92 km, and at this height electrons and positive ions diffuse at the same rate in directions transverse to the Earth's magnetic field. The asymmetry of the effective diffusion coefficient will result in elliptical cross sections for diffusing meteor trails above about 90 km. At heights above 92 km the ability of the electrons to retard the transverse motion of the more rapidly diffusing positive ions will presumably be limited, and it appears that D does not fall very much below D_i at any time. Marked departures from cylindrical symmetry are therefore not expected, and values of D derived from (3) should at most require reduction by a factor of 2 at high levels to take account of the presence of the Earth's magnetic field.

III. AMPLITUDE FLUCTUATIONS IN PERSISTENT ECHOES

According to the theory of radio reflections from meteor trails (Kaiser and Closs 1952) the short decay type of echo is characterized by a duration, defined as the time required for the echo amplitude to fall to $1/e$ of its initial value, which is independent of the line density α . The decay follows the exponential law (1) after an initial rapid rise to maximum amplitude. The long-enduring type of echo, on the other hand, shows a slower rise in amplitude to a flat maximum, and a final rapid exponential decay. Examples of these two types of echo are given in Figure 1, in which the amplitudes are plotted to logarithmic scale.

Echoes whose duration, for the Adelaide wavelength of 11.2 m, exceeds 2 sec rarely show the regular rise and decay in amplitude exemplified by the

echo shown in Figure 1. The amplitude of such a persistent echo usually fluctuates irregularly. Often it cannot be measured, as with the c.w. technique it is necessary to pick out the times when the sky (reflected) wave is in phase or anti-phase with the ground wave, and if the echo is too confused these times cannot be identified. Such complex echoes are often associated with an irregular "Doppler" period, i.e. the beat period between sky wave and ground wave; and, less commonly, with the appearance of multiple range traces whose structure

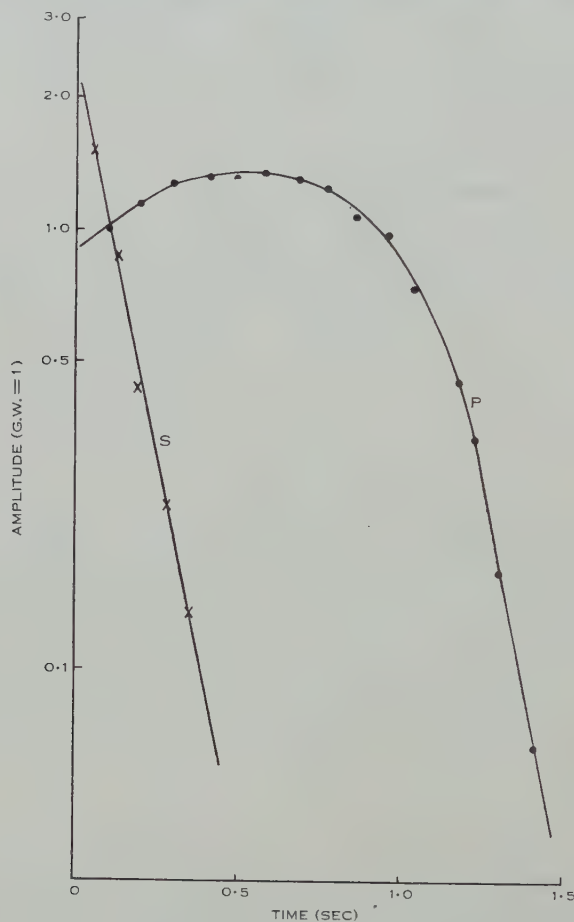


Fig. 1.—Amplitudes of typical echoes of short (curve S) and persistent (curve P) types.

may vary with time. Also, many echoes have been noted in which the amplitude fluctuations set in some time after the commencement of the echo. These facts confirm the interpretation of the irregular amplitude fluctuations in terms of the distortion of an initially uniform trail into several reflecting centres (Greenhow 1952).

Whilst it is believed that the establishment of multiple scattering centres is due to atmospheric turbulence, echoes of the type sketched in Figure 2 suggest

that the scale of the turbulence is not necessarily large. Figure 2 (a) is an example of a trail of constant slant range, reflecting primarily from two centres. The relative velocity v of these two centres is found from the formula

$$v = \lambda/2T,$$

where T is the period of the amplitude fluctuation. Since $T = 0.77$ sec for this echo, $v = 7$ m/sec. The Doppler period is remarkably constant over the whole duration of the echo, the mean half-period over 40 Doppler cycles being 0.065 ± 0.006 sec, corresponding to a line-of-sight velocity of the trail drifting in the local wind of 43 m/sec. Figure 2 (b) shows a more confused echo, but again the slant range remains constant and the mean Doppler half-period of

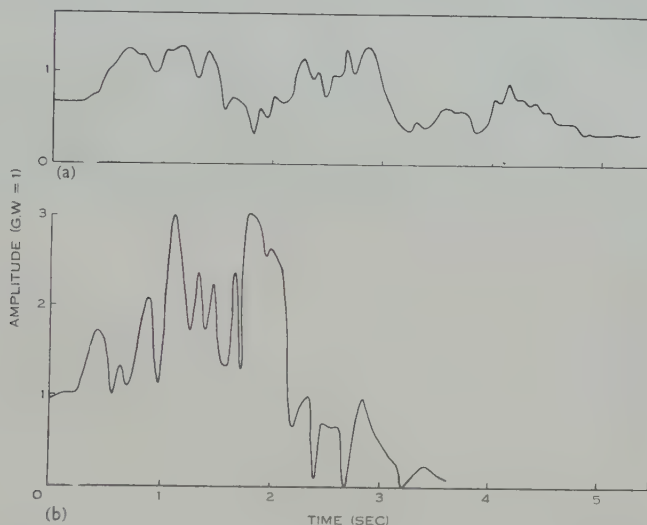


Fig. 2.—Examples of amplitude fluctuations in persistent echoes.

0.071 ± 0.009 sec does not alter over the duration of the echo. Echoes such as these, which are by no means uncommon, are presumably produced by trails broken up into two or more reflecting centres which, however, all drift in the local wind with essentially the same velocity.

Finally, a small number of persistent echoes, regular in amplitude, show very rapid slant range drifts of the order of 20 km/sec or more. According to Ellyett (1950) such rapid range drifts are caused by bending of the trail in a small wind gradient.

IV. THE MEASURED DIFFUSION COEFFICIENTS

(a) *Method of Measurement*

The rates of decay of over 1000 echoes of known heights have been determined by measurement of the amplitudes at successive maxima and minima of the Doppler beat pattern. Only those echoes which show a regular exponential decay have been used, and the rate of decay of an individual echo is found by fitting a straight line to the plot of the logarithm of the amplitude v . time.

The slope of this line fixes the time in seconds, τ , for the echo amplitude to decay to $1/e$ of its initial value, and from (1) we find for the diffusion coefficient

$$D = 0.80 \times 10^4 / \tau$$

for $\lambda = 11.2$ m. The method of height determination has been described by Robertson, Liddy, and Elford (1953).

TABLE I
MEASURED DIFFUSION COEFFICIENTS

Date	Type of Meteor	Number in Group	Mean Height (km)	Mean D ($\times 10^{-4}$)	Slope of $\ln D$ v. h ($\times 10^{-3}$)
December 10-15, 1952	Geminid	70	89.3	3.21	2.28
16-20, 1952	Sporadic	112	89.7	3.61	0.89
June 5-12, 1953	ζ -Perseid	43	90.4	4.42	1.37
	Arietid	70	89.5	4.00	0.95
	Sporadic	170	89.5	3.27	1.21
September 7-30, 1953	Sporadic	539	91.2	3.06	1.23

The echoes measured comprise both shower and sporadic meteors detected during December 1952 and June and September 1953.

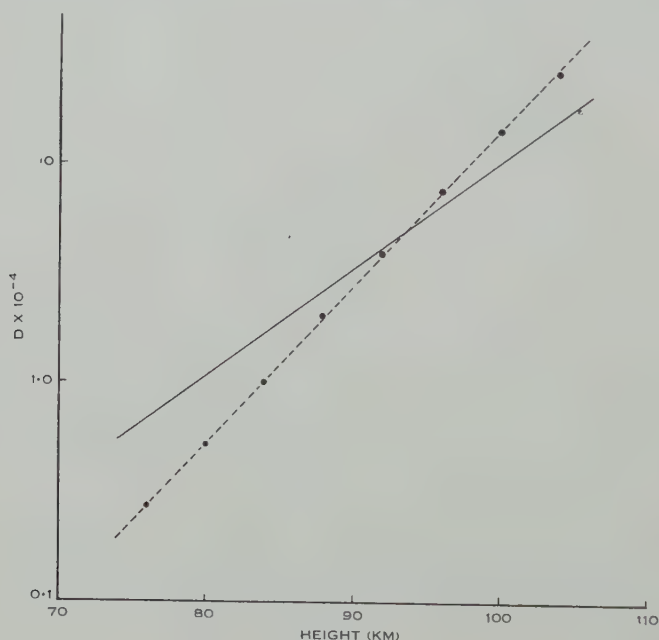


Fig. 3.—The dependence of the diffusion coefficient D upon height h .
..... Theoretical. ——— Measured.

(b) The Height Gradient

The mean height and mean diffusion coefficient for each group of echoes are given in Table 1, along with the slope of the linear relation between $\ln D$ and the height h , found by the method of least squares. The relation between D and h , based on all echoes measured, is sketched in Figure 3, together with the

theoretical relation derived in Section II. In making the least squares analysis of the measurements it has been assumed that the diffusion coefficient increases exponentially with height, which would be true only for an isothermal atmosphere. Since, however, the temperature changes in the region of interest are much smaller than the pressure changes, the expression (3) for D is dominated by the pressure p , and this assumption is not unreasonable.

It will be seen from Figure 3 that the theory of Section II predicts correct values of D . However, the measured increase of D with height is somewhat less than that predicted theoretically. A possible reason for this may lie in the selection of echoes for measurement. At heights above 95 km the decay of many echoes is so rapid that the echo amplitude falls below receiver noise level in less than one complete Doppler cycle; the rates of decay of such short echoes cannot be measured with sufficient accuracy and they are rejected. At heights below 80 km many of the more slowly decaying echoes show an irregular decay and are likewise rejected. It is therefore to be anticipated that the agreement between the theoretical and the measured height gradients is even better than is suggested by Figure 3.

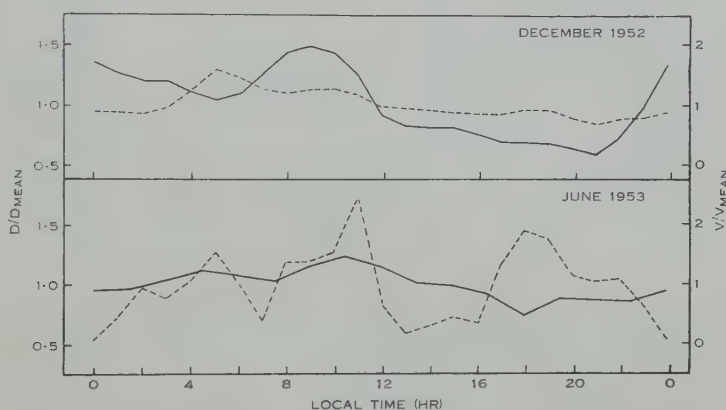


Fig. 4.—Comparison between the diurnal variations in the diffusion coefficient D and the wind speed V .

— D/D_{mean} . - - - - V/V_{mean} .

(c) The Diurnal Variation

If the dependence of D upon height is removed by dividing the data into suitable height groups and then forming D/D_{mean} , where D_{mean} is the mean value of D for a given height group, the diurnal variation of D may be studied. The diurnal variations so found for the June and December sporadic groups are shown in Figure 4. For the more numerous September sporadic group it was possible to obtain the diurnal variation in D for different height groups separately, without resort to averaging. The September variation is similar to that during June and December, with morning maxima and evening minima. The data for September also confirm that the diurnal variation is present over the whole height range of the meteor zone, with a tendency for the relative variation in D to increase with height.

Since it is possible that a meteor trail may be deformed, and its rate of decay altered, by magneto-dynamic forces which operate as the trail moves in the Earth's magnetic field, the data have been examined for correlation between the diffusion coefficient and the local wind speed. Some evidence that a high value of D is associated with a high wind speed is contained in Figure 5, which refers only to December 1952. Each point represents the average of a selected group of echoes, occurring within a period of 4 hr or less, the groups themselves being distributed from 06 to 20 hr.

But the attempt to relate the diurnal variation in D to the pattern of the local wind speed is only partially successful. During December, when the eastwards prevailing wind is particularly strong, the diurnal variation of wind

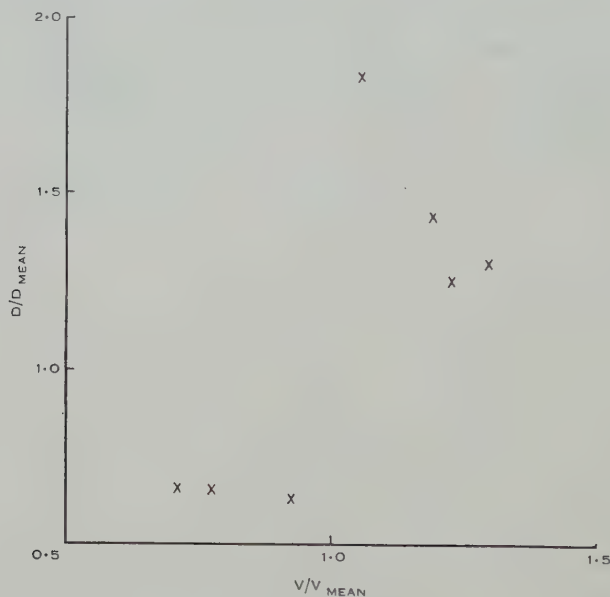


Fig. 5.—Correlation between wind speed and diffusion coefficient for selected groups of echoes, December 1952.

speed V follows closely the diurnal variation in D (see Fig. 4). For June the agreement, although not so marked, is reasonable. During September, when the wind pattern is very confused, the diurnal variation in the wind speed is almost in phase quadrature with the variation in the diffusion coefficient.

Undoubtedly, the diurnal variation in D is not wholly produced by variations in upper atmosphere winds. Although the phase of this diurnal variation appears to be incompatible with a thermal origin under solar influence, such an explanation cannot be ruled out in the present state of knowledge.

V. SCATTER AMONGST INDIVIDUAL MEASUREMENTS

Diffusion coefficients for the individual echoes show considerable scatter about the mean D v. h relation. Such scatter is evident in the mean values of D , and in the mean slopes, for the different groups listed in Table 1. It is more

clearly brought out in Figure 6, in which the individual values for the December 1952 sporadic group are plotted. Although the shower echoes give larger mean values of D and larger slopes than the sporadic echoes, the scatter between the groups is so large that it is doubtful whether this apparent difference between shower and sporadic echoes is significant.

The greater part of this unexpectedly large scatter amongst the values of D derived from individual echoes is to be ascribed to the diurnal variation already discussed. Two other processes which may contribute to the scatter are resonance effects and the influence of the Earth's magnetic field.

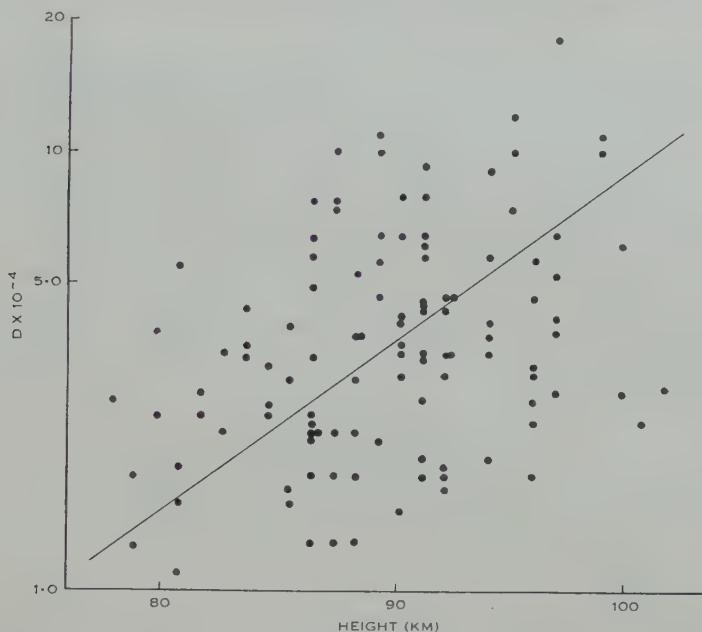


Fig. 6.—Scatter diagram of individual diffusion coefficients for 112 sporadic meteors, December 1952.

(a) Resonance Effects

The enhancement of echo amplitude in transverse, relative to parallel, scattering is confined to the initial stages of the echo, and is not expected to be large for the echoes detected by the equipment, for which the line density of electrons exceeds $10^{11}/\text{cm}$. However, the existence of such polarization effects has been demonstrated by Robertson (1953) in echoes detected at Adelaide. Whilst they may contribute to the scatter in measured diffusion coefficients, they are not considered to be of major importance. Some abnormal echoes which gave diffusion coefficients whose values decreased with time may have been influenced by plasma resonance effects.

(b) Magnetic Effects

It was suggested in Section II that the rate of diffusion of a meteor trail may be influenced by its orientation relative to the Earth's magnetic field.

The reduction in D should be greatest for trails lying parallel to the magnetic field, and is expected to become effective at heights exceeding 95 km. There is no evidence in the measured data for any general falling off of the rate of increase of D with height up to 105 km, but it is apparent that magnetic effects could produce considerable scatter in D , especially at the greater heights. Unfortunately, the number of shower echoes, for which the direction of the trail is at least roughly known, is too few to test this hypothesis: and for the more numerous sporadic echoes only the directions of the reflection points, and not the orientations of the trails, are known.

VI. CONCLUSIONS

The theory of ambipolar diffusion of meteor trails, even in its present simplified form, appears adequate to account for the broad features of the decay of meteor echoes, and in particular for the variation of the diffusion coefficient with height. It does, however, appear desirable to confirm the absolute values of the diffusion coefficient by laboratory measurements of the mobilities of the meteor ions themselves.

It is clear from the scatter amongst the individual diffusion coefficients that it is impossible to determine the height of an individual meteor trail from the rate of decay of the echo produced by it. The relation between diffusion coefficient and height can only be applied, with confidence, statistically and to very large samples.

VII. ACKNOWLEDGMENTS

The author wishes to thank Professor L. G. H. Huxley for his interest in this work and for permission to reproduce in Section II his unpublished calculation of the effective diffusion coefficient. He is also grateful to Dr. W. G. Elford for valuable discussion and for the measurement of diffusion coefficients for the Geminid echoes.

VIII. REFERENCES

- CROMPTON, R. W., HUXLEY, L. G. H., and SUTTON, D. J. (1953).—*Proc. Roy. Soc. A* **218**: 507.
 ELLYETT, C. D. (1950).—*Phil. Mag.* **41**: 694.
 HERLOFSEN, N. (1951).—*Arkiv För Fysik* **3**: 247.
 HUXLEY, L. G. H. (1952).—*Aust. J. Sci. Res. A* **5**: 10.
 GREENHOW, J. S. (1950).—*Phil. Mag.* **41**: 682.
 GREENHOW, J. S. (1952).—*Proc. Phys. Soc. B* **65**: 169.
 GREENHOW, J. S., and NEUFELD, E. L. (1955).—*J. Atmosph. Terr. Phys.* **6**: 133.
 KAISER, T. R. (1953).—*Advanc. Phys.* **2**: 495.
 KAISER, T. R., and CLOSS, R. L. (1952).—*Phil. Mag.* **43**: 1.
 MASSEY, H. S. W., and SIDA, B. W. (1955).—*Phil. Mag.* **46**: 190.
 PEARCE, A. F. (1936).—*Proc. Roy. Soc. A* **155**: 490.
 ROBERTSON, D. S. (1953).—Ph.D. Thesis, University of Adelaide.
 ROBERTSON, D. S., LIDDY, D. T., and ELFORD, W. G. (1953).—*J. Atmosph. Terr. Phys.* **4**: 255.
 ROCKET PANEL (1952).—*Phys. Rev.* **88**: 1027.
 TYNDALL, A. M. (1938).—*“The Mobility of Positive Ions in Gases.”* (Cambridge Univ. Press.)

A TEST FOR SINGULARITIES IN SYDNEY RAINFALL

By E. J. HANNAN*

[*Manuscript received November 24, 1954*]

Summary

A statistical test of the homogeneity of daily mean rainfalls (after the removal of a smooth seasonal trend) for Sydney over the years 1859 to 1952 does not contradict the hypothesis of homogeneity. The test was suggested by the theory that meteoric dust produces rainfall singularities, proposed by Bowen (1953). The test is made difficult by the complicated stochastic nature of the process generating the observations so that the significance level may in fact be lower than presumed. As the test statistic does not fall in the critical region, however, there can be no doubt that the test does not contradict the hypothesis (of homogeneity of means) at the presumed significance level.

I. INTRODUCTION

In the paper "The Influence of Meteoritic Dust on Rainfall", by E. G. Bowen (1953), it was suggested that there were certain singularities in the daily rainfall patterns of Sydney (and other cities, mainly in the southern hemisphere) and that these singularities (which represented days of exceptionally high rainfall) were due to the annually recurring deposit of dust in the atmosphere from the passage of meteor showers.

A statistical test of the theory can be obtained in two ways. One way is to relate the days of high meteor activity to the days of high rainfall, these last having been objectively chosen. The second method is to test the homogeneity of the daily means, for a station or a number of stations, after the means have been adjusted for any smooth seasonal variation. In this paper the second method will be applied to the Sydney rainfall. In the process of preparing the data for this test a sound basis for objectively choosing the days of high rainfall will be obtained, so that the first method may later be used.

II. THE NATURE OF THE DATA AND OF THE NULL HYPOTHESIS

The distribution of daily rainfall is J-shaped. In an accompanying paper (Das 1955) it has been shown that, for a period of 22 days in October and November, a type III distribution fits very well. This best-fitting distribution is so extremely skew that observations more than 16.5 standard deviations away from the mean occur with a probability of 10^{-4} . If the distribution were normal with the same mean and standard deviation the occurrence of an observation more than 6.1 standard deviation units from the mean would have the probability 10^{-9} .

* Australian National University, Canberra, A.C.T.

An indication of the very great departure from normality is also obtained by considering the daily variances. For example, over half of the January and February variances are "significantly" different from what is clearly a well-fitted trend, when treated as being distributed as χ^2 with 94 degrees of freedom. The explanation for this phenomenon is of course found in the high fourth moment of the distribution (see Box 1953).

This non-normality is the first of a number of difficulties which affect a test of the significance of the rainfall singularities. A second is the lack of independence between rainfalls on days near each other in time. A final difficulty is the seasonal variation in the data which affects not only the daily means and variances but also the degree of dependence between days.

A study of the correlations for lags of up to 4 days suggests that

$$\rho_{kj} \approx \rho_{1j}^k, \quad k=1, 2, 3, 4.$$

Here ρ_{kj} is the correlation between rainfall on days j and $j-k$. For $k > 4$ the correlation appears to be, effectively, zero. For example the first four lag correlations were computed from the observations over the 94 years between the dates October 17 and November 7 (2068 observations in all). Over this period the seasonal variation is small. The correlations were computed using the formulae

$$r_k = \frac{\sum_{i=1}^{22} \sum_{j=k+1}^{94} (x_{ij}x_{ij-k}) - 94(22-k)\bar{x}^2}{22-k \left[\sum_{i=1}^{94} \sum_{j=1}^{22} (x_{ij}^2) - 94(22)\bar{x}^2 \right]}. \quad (k=1, 2, 3, 4)$$

Here

$$\bar{x} = \frac{1}{2068} \sum_{i=1}^{94} \sum_{j=1}^{22} x_{ij},$$

and x_{ij} is the rainfall on the j th day of the i th year.

This particular formula was used mainly for computational convenience, the $\sum_{i=1}^{94} x_{ij}x_{ij-k}$ and $\sum_{i=1}^{94} x_{ij}^2$ having already been computed on the C.S.I.R.O. electronic computer so that the seasonal variation in the means, variances, and lag correlations could be estimated. The statistics r_k will provide consistent estimators of the corresponding ρ_k . A formula such as $(22-k)^{-1} \sum_{j=k}^{22} r_{kj}$, where r_{kj} is the observed correlation between days j and $j-k$, would not be satisfactory since the bias will remain the same as the bias of the individual r_{kj} . The statistics r_k , together with r_1^k , are given below.

$$r_1 = 0.219; \quad \begin{cases} r_2 = 0.057 \\ r_1^2 = 0.048 \end{cases}; \quad \begin{cases} r_3 = -0.007 \\ r_1^3 = 0.011 \end{cases}; \quad \begin{cases} r_4 = 0.048 \\ r_1^4 = 0.002 \end{cases}$$

If these were computed from normally distributed observations their standard deviations would be about 0.02. The variance in the correlation coefficient depends upon the moments of the fourth order (Cramer 1946, p. 359)

and in the present case is likely to be underestimated by the formula, $(1 - \rho^2)^{2/n}$, which applies for a normal distribution.

The high value of r_4 (compared with r_1^4) is entirely due to one cross product, that between the rainfall of 181 points on October 25, 1882 and 423 points on October 29, 1882, which is probably fortuitous and a result of the extraordinary skew parent distribution.

It is interesting to note that the serial correlations r_k in the rainfall data are here estimated by ordinary correlations (space averages), the data for the several years being treated as realizations of a stochastic process. This is a rather extraordinary situation for usually only one realization is available, from which the serial correlation is estimated as a phase average.

The relation $\rho_{kj} = \rho_{1j}^k$, which holds reasonably well for k small, at first sight suggests that the process can be represented by a model of the form

$$(x_j - m) = \rho(x_{j-1} - m) + \varepsilon_j, \quad |\rho| < 1.$$

Here m is the mean of the process, ρ the correlation between x_j and x_{j-1} , and ε_j is an independent random process with zero mean and variance $\sigma^2(1 - \rho^2)$.

It is at once evident that this process cannot be used to describe rainfall data for not only is there a seasonal variation in ρ , m , and σ^2 but the variate x_j is necessarily positive so that $\varepsilon_j + m(1 - \rho)$ would have to have a distribution admitting only positive values. This would, however, imply a smoothness in the series of x_j which would not be supported by the data. (A high x_j could not be followed by an x_{j+1} near to zero.)

A modified process which suggests itself is

$$\left(\frac{x_j - m_j}{\sigma_j}\right) = \rho_{1j} \left(\frac{x_{j-1} - m_{j-1}}{\sigma_{j-1}}\right) + \varepsilon_j, \quad \dots \dots \dots (1)$$

Here m_j is the mean and σ_j^2 the variance of x_j , ρ_{1j} is the correlation between x_j and x_{j-1} , and ε_j is an independent random process with zero mean and variance $(1 - \rho_{1j}^2)$.

It can be seen that the correlation between x_j and x_{j-k} will be

$$\rho_{kj} = \rho_{1j} \rho_{1j-1} \dots \rho_{1j-k+1}.$$

If the lag k is not too large and the seasonal effect is small, over short periods, then

$$\rho_{kj} \approx \rho_{1j}^k.$$

Again this process will not be satisfactory for the daily rainfalls (since they are positive) but it may provide a good approximation to the process generating the daily mean rainfall (over 94 years) since the range of variation of the sample mean below the true mean will be much greater. Moreover, if the m_j , σ_j^2 , and ρ_{1j} are equated to the means, variances, and first serial correlations of daily rainfall for day j (as estimated from a graduation of the observed values) the process (1) will have, approximately, the same means, variances, and serial correlations (for all lags) as the actual series.

A better approximation to the process generating the daily means could be got by including a term $(x_{j-2} - m_{j-2})/\sigma_{j-2}$. However, the nearness of the r_{2j} to r_{1j}^2 suggests that the coefficient of this term will be very small, for this coefficient will be, approximately, $(r_{2j} - r_{1j}^2)(1 - r_{1j}^2)^{-1}$. The additional computations involved would not be justified by the very small effect from the inclusion of the extra term.

The model of the process generating the daily means which will be used is, therefore,

$$\sqrt{94} \frac{\bar{x}_j - m_j}{\{\sigma_j^2(1 - \rho_{1j}^2)\}^{\frac{1}{2}}} = \rho_{1j} \sqrt{94} \frac{\bar{x}_{j-1} - m_{j-1}}{\{\sigma_{j-1}^2(1 - \rho_{1j}^2)\}^{\frac{1}{2}}} + \varepsilon_j. \quad \dots\dots (2)$$

Here \bar{x}_j is the mean rainfall observed on the j th day of the year while m_j , σ_j^2 , and ρ_{1j} are the true means, variances, and first serial correlations of rainfall on that day. The ε_j come from a process with zero mean and unit variance.

The specification of the null hypothesis will be completed if the nature of the seasonal variation in the means, variances, and correlation coefficients is laid down. All that needs to be said here is that these seasonal variations should be described by reasonably smooth curves such as a low order polynomial in j or a trigonometric polynomial formed from the first few harmonics. The choice among these alternatives depends on the goodness of their fit to the data, with the proviso that the fit should not be carried anywhere near to the point where individual singularities, confined to a period of a relatively few days, would be eliminated.

This restriction on the order of the polynomials means that a very large number of degrees of freedom are available for the estimation of the relatively few constants involved so that the m_j , $\hat{\sigma}_j^2$, and $\hat{\rho}_{1j}$ given by the estimated curves can be treated as the true values to a satisfactory degree of approximation and, when used to transform the \bar{x}_j into the corresponding $\hat{\varepsilon}_j$, these can be regarded as uncorrelated random variates with zero mean and unit variance. Since the means \bar{x}_j should have a distribution near to normality the sum $\chi_1^2 = \sum \hat{\varepsilon}_j^2$ should be approximately distributed as χ^2 with $(365 - p)$ degrees of freedom, where p is the number of constants fitted in the process of estimating the seasonal variation of the m_j , σ_j^2 , and ρ_{1j} .

This test is the generalization of the classical test of the homogeneity of a set of means (from observations subject to different treatments) which is obtained by comparing the variance within treatments with that between the means themselves. In the present case the comparison is being made between the within-days variance of rainfall and the variance as estimated from the daily means. In addition, however, the data have had to be transformed to independence by the use of the ρ_{1j} and the within-days variances made homogeneous while the effect of the seasonal variation in the means has been removed so that it will not void the test against the effect of the meteor showers.

This test is subject to a number of qualifications, however, as has already been indicated:

(a) In fact the seasonal pattern of the m_j , σ_j^2 , and ρ_{1j} will not be known exactly, but will have been estimated, so that the variance of the statistic χ_1^2 will be increased by a component due to the variance of these estimates.

(b) The true nature of the underlying process generating the x_j will be more complicated than is indicated by (2) above, so that the residuals $\hat{\varepsilon}_j$ will not be independent. Their correlations may also be different from zero. This may increase the variance of χ_1^2 also. For the variance of this quantity will be

$$E\{\sum \hat{\varepsilon}_i^2\}^2 - \{E(\sum \hat{\varepsilon}_i^2)\}^2,$$

where E denotes expected value.

This expectation will include such terms as

$$E\{\hat{\varepsilon}_i^2 \hat{\varepsilon}_j^2\}, \quad j \neq i.$$

If the $\hat{\varepsilon}_i$ were truly independent with zero mean and unit variance this expectation would be 1. If $\hat{\varepsilon}_i$ and $\hat{\varepsilon}_j$ are not independent the expectation may be different from 1. The high fourth moment of the daily rainfalls suggests that the second moment of the individual ε_i^2 will be high and that $E\{\hat{\varepsilon}_i^2 \hat{\varepsilon}_j^2\}$ may be greater than 1. This will increase the variance of χ_1^2 above its theoretical value.

(c) The extreme non-normality of the daily rainfall distribution will also increase the variance of χ_1^2 . Since χ_1^2 is itself a sum of squares of standardized observations, it is clear that its variance will depend upon the fourth moment of these observations. If this fourth moment is greater than that of a standard normal variate the variance will be greater than it would be if it were exactly distributed as χ_{365-p}^2 . In the present case the nature of the parent distribution makes it certain that the fourth moment will be high. The effect will, of course, be reduced by the fact that each $\hat{\varepsilon}_j$ is the sum of a large number of observations but, because of the very extreme parent distribution, it will probably persist.

(d) Since the large deviations in the present distribution are all positive the distribution of χ_1^2 will also have a greater positive skewness than it should theoretically have.

The effect of all of these factors will therefore be to increase the variance of χ_1^2 above its theoretical value. The fourth factor, which with the third seems to be the most important, will also make the distribution more skew (with a longer tail to the right) than it should have. The test of significance will consist of choosing a positive number, t , such that the probability of an observed χ_1^2 being greater than or equal to t (and therefore significant) is α . Both the increase in variance and the increase in positive skewness will make the probability of a significantly large χ_1^2 occurring (if the null hypothesis is true) greater than α .

If, therefore, the observed χ_1^2 is just beyond the critical point it will not follow that the result is significant at the level α since the true critical point may be further to the right by a sufficient amount to render this conclusion invalid. On the other hand, if the observed χ_1^2 is to the left of the critical point it can be said that the result is not significant and the null hypothesis will not be rejected at the level α .

This element of indeterminacy is of course a defect in the test. It is one which cannot easily be removed, however. Even if an exact distribution-free test, based, for example, on rearrangements of the observations, could be found it would almost certainly be much less powerful than the present test.

III. REMOVAL OF SEASONAL VARIATION

In order to get a closer agreement between calendar and sidereal years the leap year was reinserted in 1900 so that March 1, 1900 became February 29, 1899, and so on.

Since a strictly periodic seasonal effect was being removed it seemed clear that a harmonic curve should be used to graduate the series. At the same time no attempt was made to fit a harmonic curve including terms with too short a period, as this would tend to remove the effect being sought.

For both means and variances the year was divided into 24 periods of equal lengths and the first four harmonics were fitted to the resulting 24 values. The intensities S^2 of these four harmonics, expressed as multiples of $4V^2/24$ are shown in Table 1. (Here V^2 is the estimate of the variance of the 24 values.)

The intensity corresponding to the period μ is derived from the relation

$$S^2 = A^2 + B^2,$$

$$A^2 = \frac{2}{24} \sum_j u_j \cos \frac{2\pi j}{\mu}, \quad B^2 = \frac{2}{24} \sum_j u_j \sin \frac{2\pi j}{\mu},$$

where u_j is the mean (or variance) on the j th day.

TABLE I
 $\kappa = 24S^2/4V^2$

Period (days)	Means	Variances
365	8.53	9.57
$\frac{365}{2}$	0.09	0.06
$\frac{365}{3}$	0.17	0.10
$\frac{365}{4}$	0.13	0.24

There is, of course, no doubt of the significance of the period of 365 days. On the other hand the quantities κ for the other three periods are not significant at any reasonable level. It was also clear from an examination of the data that little could be gained by adding further terms unless their periods became quite short.

The first harmonic was therefore fitted to the means and variances using the individual daily averages. The resulting curves were:

$$\text{Means:} \quad \hat{m}_j = 12.865 + 4.377 \sin \left(\frac{2\pi j}{365.25} + 0.527 \right)$$

$$\text{Variances:} \quad \sigma_j^2 = 1755.0 + 1015.9 \sin \left(\frac{2\pi j}{365.25} + 0.571 \right)$$

Here $j=1$ on March 1. The phase angles do not differ significantly.

A simple harmonic was also fitted to the between-days correlations, the resulting curve being

$$\hat{\rho}_{1j} = 0.338 + 0.103 \sin \left(\frac{2\pi j}{365 \cdot 25} - 0.285 \right).$$

Again $j=1$ on March 1.

The fit of this curve is not as good as that to the means and variances and it seems, in retrospect, that higher order terms should have been taken. In fact if the first harmonic were fitted to the covariances rather than the correlations a better fit would be obtained and it seems that the phase angle would not differ significantly from those for the means and variances. The effect on the residuals of the poor fit to the r_{1j} will be small, however (it will increase the variance of χ_1^2).

IV. TEST OF HOMOGENEITY OF MEANS

The residuals

$$\hat{\varepsilon}_j = \left(\frac{94}{1 - \hat{\rho}_{1j}^2} \right)^{\frac{1}{2}} \left\{ \frac{\bar{x}_j - \hat{m}_j}{\hat{\sigma}_j} - \hat{\rho}_{1j} \frac{\bar{x}_{j-1} - \hat{m}_{j-1}}{\hat{\sigma}_{j-1}} \right\} \dots \dots \dots (3)$$

were formed. The $\hat{\varepsilon}_j$ had a small positive mean (0.018) and a small negative serial correlation (-0.05), both of which were far from significant.

The distribution of the $\hat{\varepsilon}_j$ approaches normality, as could be expected, but there is positive skewness.

There are 366 $\hat{\varepsilon}_j$, one resulting from the 24 rainfalls on February 29. (For this $\hat{\varepsilon}_j$ the factor $(24)^{\frac{1}{2}}$ will replace $(94)^{\frac{1}{2}}$ in (3).) Three constants have been fitted to each of the means, variances, and serial correlations so that χ_1^2 may be treated as distributed as χ^2 with 357 degrees of freedom. The value of $\Sigma \hat{\varepsilon}_j^2 = \chi_1^2$ is 394.1. Using Fisher's approximation to the distribution of χ^2 the quantity

$$\{2(394.1)\}^{\frac{1}{2}} - \{2(357) - 1\}^{\frac{1}{2}} = 1.37$$

was computed. Treated as a standard normal variate and using one tail of the normal distribution this is well inside the 5 per cent. point (1.65).

The null hypothesis, therefore, cannot be rejected at this level of significance and the result of the test, while not of course disproving Bowen's hypothesis, cannot be said to justify it.

V. ALTERNATIVE TESTS

The test which has been given might not be very powerful against the alternative of only a few very widely scattered singularities since the effect of these in increasing χ_1^2 may be lost among the accompanying "noise".

At first sight an alternative would be to pick out the high peaks in the residuals and test them as the largest among 366. For example, the highest peak in the residuals is on July 23 (24, 1901-1952). This residual is 4.037 times its (trend) standard deviation. This lies almost exactly on the 1 per cent. point for the largest out of 366 normally and independently distributed observations. In computing the probability of obtaining a *largest* observation, as

extreme as or more extreme than the largest observed, only the extreme tail area of the distribution of the individual observations is used. Moreover any error here will be multiplied, in the process of computing the first probability, by the number of observations (approximately). It is certain, from the shape of the parent distribution, that the tail area used will be too small compared with the true, unknown, probability. How badly misleading the assumption of normality can be here can be seen by considering the distribution of χ^2 for 102 degrees of freedom. (The largest number of degrees of freedom available in Pearson's tables of the "Incomplete Gamma Function".) This distribution, to the eye, would be indistinguishable from the normal distribution. In the

TABLE 2
DAYS FOR WHICH THE RESIDUALS $\hat{\epsilon}_j$ ARE INDIVIDUALLY SIGNIFICANT AND POSITIVE

Date (prior to 1900)	Total Rainfall on this Date over 94 Years (points)	Total of Four Highest Rain- falls over 94 Years (points)	$\frac{\text{Col. 3}}{\text{Col. 2}} \times 100$
Jan. 11.. ..	1792	1170	65
12.. ..	2293	1296	57
Mar. 20.. ..	2603	1307	50
Apr. 6.. ..	3293	1943	59
May 6.. ..	2382	1159	49
25.. ..	2758	1412	51
July 23.. ..	2826	1043	37
Aug. 2.. ..	1866	1148	62
23.. ..	1635	613	37
Sept. 24.. ..	1512	836	55
28.. ..	1754	1067	61
Oct. 12.. ..	1452	1066	73
29.. ..	1186	743	63
Nov. 17.. ..	1315	540	41
18.. ..	1897	871	46
Dec. 1.. ..	1505	905	60

accompanying paper Das has shown that for the period from October 17 to November 7 the distribution of the daily means is close to χ^2 with 10 degrees of freedom. The distribution of the residuals $\hat{\epsilon}_j$ is therefore much further from normal than χ^2_{102} . The 5 per cent. point of the greatest out of 366 observations from χ^2_{102} is 4.2 standard deviation units from the mean. The corresponding 5 per cent. point for the normal distribution is 3.6 standard deviation units from the mean. The largest observed rainfall residual, while significant at the 1 per cent. point on the basis of normality, is not significant at the 5 per cent. point on the basis of χ^2_{102} . This result suggests that true probability of a largest residual ≥ 4.037 is much higher than 0.05. The large effect on the computed probabilities of small variations in the nature of the parent distribution certainly makes the testing of the significance of the large observations an impossible task.

Further and more powerful tests can perhaps be obtained by a comparison of the sequence of meteor showers throughout the year with the days of high residuals or by considering data for a number of different stations.

It is of some interest to examine the days for which the residuals $\hat{\varepsilon}_j$ are individually "significant" (and positive), say at the 1 per cent. point. These days are listed in Table 2 (under their dates prior to 1900), which also shows the total rainfall on each of these days over the 94 years and the total rainfall on the 4 occasions of highest rainfall.

The highest rainfall over the years for these 16 days is, on an average, 22.7 per cent. of the total rainfall. Most or all of the deviation from the mean seems therefore to be due to a relatively few very heavy falls of rain. This phenomenon is, to some extent at least, explained by the very skew distribution of daily rainfall.

VI. ACKNOWLEDGMENTS

I wish to thank Dr. E. G. Bowen for suggesting this investigation, and Mr. Pearcey and Mr. Beard of C.S.I.R.O. for the assistance they gave with the C.S.I.R.O. electronic computer.

VII. REFERENCES

- BOWEN, E. G. (1953).—*Aust. J. Phys.* **6**: 490.
BOX, G. E. P. (1953).—*Biometrika* **40**: 318.
CRAMER, H. (1946).—"Mathematical Methods of Statistics." (Princeton Univ. Press.)
DAS, S. C. (1955).—*Aust. J. Phys.* **8**: 298.

THE FITTING OF TRUNCATED TYPE III CURVES TO DAILY RAINFALL DATA

By S. C. DAS*

[*Manuscript received November 24, 1954*]

Summary

The method of maximum likelihood has been used to fit a truncated type III(Gamma) distribution to daily rainfall data for Sydney over the period 1859–1952. An approximate test of the hypothesis that there is a singularity at the origin is suggested. This test is based on a comparison of the expected frequency in the truncated part, when the observed frequency in this part is taken into account in the fit, with the expected frequency when these observations are neglected. For Sydney the test shows that there is no evidence in the rainfall data for a singularity at the origin.

I. INTRODUCTION

This paper discusses a problem of curve fitting which arose in testing the hypothesis proposed by Bowen (1953) concerning daily rainfall data. The hypothesis advanced is that meteoritic dust is an important factor in stimulating rainfall.

An analysis of the data to test this hypothesis is presented by Hannan simultaneously with this paper (Hannan 1955). In connexion with this, it is of interest to fit a frequency distribution to the daily rainfall data in order to judge the effect of departure from normality in the distribution of daily rainfall on the test of Bowen's hypothesis.

The rainfall figures (Sydney 1859–1952) which constitute the data for the curve fitting refer to a period of 22 days from October 17 to November 7 for 94 years. The reason for choosing the period between October and November was the small seasonal variation during this period : a period of 22 days was taken so that the total observations might exceed 2000. In fact there are 2068 observations. The shape of the distribution of rainfall suggests that a type III probability distribution of the form :

$$f(x) = \frac{\mu^\chi}{\Gamma(\chi)} e^{-\mu x} x^{\chi-1}$$

might provide a good fit. There are, however, a large number of zero observations. This makes it impossible to apply the ordinary maximum likelihood equations which involve the sum of logarithms of the observations. We have therefore modified the ordinary method by truncating the curve, and applying a modified maximum likelihood method which takes account of the number of observations in the truncated part. The resulting theory and the calculations

* Australian National University, Canberra, A.C.T.

for the particular example are given in Section II. A very good fit is obtained as judged by the χ^2 test, and, in particular, there is a close agreement between the expected and observed numbers in the truncated part.

One might, at first sight, have suspected that such a good fit would not have been found and that the best way to fit observations of this kind would be to fit a mixed probability distribution which had a non-zero concentration at zero, and a continuous distribution for values not equal to zero. This is clearly not so in the present case, but a method is suggested for testing such an alternative. To do this we fit a truncated type III curve to the observations, ignoring the numbers in the truncated part, and compare the observed numbers in the truncated part with the numbers expected from the fitted curve. As an illustration of this method this is done in Section III.

In this connexion it is worth mentioning work done previously in this direction. Cohen (1950) uses the method of moments to develop formulas for testing μ , σ , α_3 , the population mean, standard deviation, and the third standard moment respectively from a singly truncated sample when the population is distributed according to Pearson's type III function. Des Raj (1953) discusses the theory of estimation for the same population parameters as Cohen for a type III curve from both singly and doubly truncated samples. He obtains estimating equations by the method of moments and has also shown that they can be obtained by the method of maximum likelihood.

In our present problem we know where the origin of the parent distribution is whereas Cohen and Des Raj are trying to estimate it and so are estimating three parameters. Here we have two parameters and we have evolved a method suitable for our problem which is different from both of theirs.

II. FITTING OF TRUNCATED CURVE TAKING THE NUMBER OF OBSERVATIONS IN THE TRUNCATED PART INTO ACCOUNT

In this section we first discuss the difficulty that is to be encountered in estimating the parameters on account of zero values. Let x_1, x_2, \dots, x_N be a random sample of size N from a type III distribution given by

$$f(x) = \frac{\mu^\kappa}{\Gamma(\kappa)} e^{-\mu x} x^{\kappa-1}, \quad \dots \dots \dots (1)$$

The likelihood of the observations is given by

$$\varphi(x_1, x_2, \dots, x_N) = \frac{\mu^{N\kappa}}{\{\Gamma(\kappa)\}^N} \exp(-\mu \sum_{i=1}^N x_i) \prod_{i=1}^N x_i^{\kappa-1}, \quad \dots \dots (2)$$

Taking the logarithm of (2) we get

$$L = \ln \varphi\{x_1, x_2, \dots, x_N\} = N\kappa \ln \mu - N \ln \Gamma(\kappa) - \mu \sum_{i=1}^N x_i + (\kappa-1) \sum_{i=1}^N \ln x_i, \quad \dots \dots \dots (3)$$

Thus for estimating the parameters μ and κ we have the following maximum likelihood equations :

$$\frac{1}{N} \frac{\partial L}{\partial \mu} = \frac{\kappa}{\mu} - \frac{1}{N} \sum_{i=1}^N x_i = 0, \quad \dots \dots \dots (4)$$

$$\frac{1}{N} \frac{\partial L}{\partial \kappa} = \ln \mu - \frac{d \ln \Gamma(\kappa)}{d\kappa} + \frac{1}{N} \sum_{i=1}^N \ln x_i = 0. \quad \dots \dots \dots (5)$$

Now in practice the fact that we measure the x_i to the nearest rounded-off unit on some scale means that in many cases there will be zero values of x_i , and in fact in the case of the daily rainfall there are a large number of zero values. When this happens equation (5) cannot be used.

To avoid this difficulty we choose a small interval $(0, \delta)$ and truncate the distribution at δ . We ignore the actual values of x_i less than δ , but use the fact that we know their total number. Thus if n be the number of observations falling in $(0, \delta)$, the rest $N - n = m$ of the observations will all be greater than δ .

The likelihood function in this case is given by

$$\begin{aligned} \varphi(x_1, \dots, x_m) &= \binom{N}{n} \frac{\mu^{n\kappa}}{\{\Gamma(\kappa)\}^n} \left\{ \int_0^\delta \exp(-\mu x) x^{\kappa-1} dx \right\}^n \frac{\mu^{m\kappa}}{\{\Gamma(\kappa)\}^m} \\ &\times \exp\left(-\mu \sum_{i=1}^m x_i\right) \prod_{i=1}^m x_i^{\kappa-1}, \quad \dots \dots \dots (6) \end{aligned}$$

where $x_1, x_2, \dots, x_m \geq \delta$. From this we get

$$\begin{aligned} L = \ln \varphi &= \ln \binom{N}{n} + N\kappa \ln \mu - N \ln \Gamma(\kappa) + n \ln \int_0^\delta e^{-\mu x} x^{\kappa-1} dx \\ &- \mu \sum_{i=1}^m x_i + (\kappa - 1) \sum_{i=1}^m \ln x_i. \quad \dots \dots \dots (7) \end{aligned}$$

Now when $\delta \rightarrow 0$

$$\int_0^\delta e^{-\mu x} x^{\kappa-1} dx \sim \int_0^\delta x^{\kappa-1} dx = \frac{\delta^\kappa}{\kappa},$$

and substituting this value in (7) we get

$$\begin{aligned} L &= \ln \binom{N}{n} + N\kappa \ln \mu - N \ln \Gamma(\kappa) + n \ln \delta - n \ln \kappa - \mu \sum_{i=1}^m x_i + (\kappa - 1) \sum_{i=1}^m \ln x_i \\ &\dots \dots \dots (8) \end{aligned}$$

The maximum likelihood equations are consequently given by

$$\frac{1}{N} \frac{\partial L}{\partial \mu} = \frac{\kappa}{\mu} - \frac{1}{N} \sum_{i=1}^m x_i = 0, \quad \dots \dots \dots (9)$$

$$\frac{1}{N} \frac{\partial L}{\partial \kappa} = \ln \mu - \frac{d \ln \Gamma(\kappa)}{d\kappa} + \frac{n}{N} \ln \delta - \frac{n}{N\kappa} + \frac{1}{N} \sum_{i=1}^m \ln x_i = 0. \quad \dots \dots \dots (10)$$

For some values of α and μ the approximation resulting from replacing

$$\int_0^{\delta} e^{-\mu x} x^{\alpha-1} dx \text{ by } \frac{\delta^{\alpha}}{\alpha}$$

may not be good enough and then the solution of the equations is a little more laborious.

It is also easy to show that in most cases the loss of information arising because we do not have exact values x_i and consequently have to use (6) instead of (2), is quite small.

As a numerical example we give in Table 1 Sydney rainfall values for 2068 days.

TABLE 1
FREQUENCIES AS PREDICTED BY THE FITTED CURVE

Class Interval	f_0	$f_E^{(1)}$	$f_E^{(2)}$
0-5	1631	1638.5	(1614.0)
6-10	115	106.0	103.6
11-15	67	62.0	62.2
16-20	42	43.6	43.6
21-25	27	32.2	32.9
26-30	26	26.0	26.0
31-35	19	20.7	21.1
36-40	14	17.2	17.5
41-45	12	14.3	14.6
46-50	18	12.2	12.5
51-60	18	19.6	20.2
61-70	13	14.7	15.4
71-80	13	11.6	12.0
81-90	8	8.9	9.6
91-100	8	7.2	7.6
101-125	16	12.2	13.5
125-150	7	7.2	8.3
150-425	14	13.2	15.7

Now for estimating μ and α from equations (9) and (10) we take $\delta=5$ and consequently we have $n=1631$,

$$\sum_{i=1}^m x_i = 16891 \quad \sum_{i=1}^m \ln x_i = 1373.117.$$

Substituting these values in (9) and (10) we find that equation (9) reduces to $\mu=0.1224\alpha$ and equation (10) reduces to

$$\frac{d \ln \Gamma(\alpha)}{d\alpha} \ln \alpha - \frac{0.2114}{\alpha} + 0.1703 = 0.$$

Solving these equations we get $\alpha=0.105$, $\mu=0.013$ correct to three places of decimals.

Now for these values of μ , κ , and δ we have

$$\int_0^{\delta} e^{-\mu x x^{\kappa-1}} dx = 11.14 \quad \text{and} \quad \frac{\delta^{\kappa}}{\kappa} = 11.30.$$

The calculation of the expected frequencies for testing the goodness of fit is shown in the table. Since the χ^2 test is used for testing goodness of fit, the observations are grouped into classes so that the expected frequency in any class is not less than 5. To calculate the expected frequencies we use tables of the incomplete Γ -function (Pearson 1922) making a double linear interpolation, which is sufficiently accurate for our purpose.

The results are shown in Table 1. The first column gives the class interval, the column headed f_0 gives the corresponding observed class frequencies. The column headed $f_E^{(1)}$ gives the expected frequencies. Thus for 15 degrees of freedom the total χ^2 is found to be 7.8, which shows that the fit is an extremely good one.

III. FITTING OF TRUNCATED CURVE IGNORING THE OBSERVATIONS IN THE TRUNCATED PART

Here we fit a truncated type III curve to the observations which are all greater than δ and ignore all observations which are less than δ . The probability density in this case is given by

$$f(x) = \frac{\mu^{\kappa}}{\Gamma(\kappa)} \cdot \frac{e^{-\mu x x^{\kappa-1}}}{1 - \frac{\mu^{\kappa}}{\Gamma(\kappa)} \int_0^{\delta} e^{-\mu x x^{\kappa-1}} dx}.$$

The logarithm of the likelihood function is given by

$$L = \ln \varphi(x_1, \dots, x_m) = m\kappa \ln \mu - m \ln \Gamma(\kappa) - mG(\mu, \kappa) - \mu \sum_{i=1}^m x_i + (\kappa - 1) \sum_{i=1}^m \ln x_i, \quad \dots \dots \dots (11)$$

where

$$G(\mu, \kappa) = \ln \left\{ 1 - \frac{\mu^{\kappa}}{\Gamma(\kappa)} \int_0^{\delta} e^{-\mu x x^{\kappa-1}} dx \right\}.$$

From (11) we obtain the following likelihood equations:

$$\frac{1}{m} \frac{\partial L}{\partial \mu} = \frac{\kappa}{\mu} - \frac{\partial G}{\partial \mu} - \frac{\sum x_i}{m} = 0, \quad \dots \dots \dots (12)$$

$$\frac{1}{m} \frac{\partial L}{\partial \kappa} = \ln \mu - \frac{d \ln \Gamma(\kappa)}{d \kappa} - \frac{\partial G}{\partial \kappa} + \frac{\sum \ln x_i}{m} = 0. \quad \dots \dots \dots (13)$$

For these equations δ , $\sum x_i$, and $\sum \ln x_i$ have the same values as before and $m=437$. Equations (12) and (13) are generally too complicated to be solved. But we can, however, find approximate solutions and then improve these solutions to any desired degree of accuracy. Thus if $\mu = \mu_0$ and $\kappa = \kappa_0$ are an approximate solution for the equations (12) and (13) a better approximate

solution is obtained by taking $\mu = \mu_0 + \delta\mu_0$, $\kappa = \kappa_0 + \delta\kappa_0$, where $\delta\mu_0$, $\delta\kappa_0$ are given by the following equations :

$$\left(\frac{\partial^2 L}{\partial \mu^2}\right)_{\mu_0} \delta\mu_0 + \left(\frac{\partial^2 L}{\partial \mu \partial \kappa}\right)_{\mu_0} \delta\kappa_0 = -\left(\frac{\partial L}{\partial \mu}\right)_{\mu_0} \quad \dots\dots\dots (14)$$

$$\left(\frac{\partial^2 L}{\partial \mu \partial \kappa}\right)_{\mu_0} \delta\mu_0 + \left(\frac{\partial^2 L}{\partial \kappa^2}\right)_{\mu_0} \delta\kappa_0 = -\left(\frac{\partial L}{\partial \kappa}\right)_{\mu_0} \quad \dots\dots\dots (15)$$

In this particular example we took $\kappa = 0.10$ and $\mu = 0.01$ as the approximate solution for equations (12) and (13) : then with the help of equations (14) and (15) we found $\kappa = 0.105$ and $\mu = 0.012$ correct to three places of decimal. The Newton-Gregory formula was used for numerical differentiation in calculating the $\partial G/\partial \mu$, $\partial^2 G/\partial \mu^2$, $\partial G/\partial \kappa$, $\partial^2 G/\partial \kappa^2$, and $\partial^2 G/\partial \mu \partial \kappa$, etc., which occur in the solutions of equations (14) and (15). In this case, the approximation previously given for the integral in G is not sufficiently accurate. The expected frequencies are given in Table 1 in the column headed $f_E^{(2)}$. The fit of the observed frequencies to the expected values, ignoring the interval $(0, \delta)$ gives a $\chi^2 = 8.47$ which shows that the fit is a good one.

The expected frequency in the interval $(0, \delta)$ as predicted by this fitted distribution is 1614. To test whether this is significantly different from the observed frequency in this interval we use the following statistic :

$$E = \frac{N_1 - Np}{\sqrt{Npq + V(Np)}},$$

where N_1 stands for the number observed in the range $(0, \delta)$, N stands for the number of observations in the whole sample, and p is the probability of an observation falling in the range $(0, \delta)$ given by

$$p = \frac{\mu^\kappa}{\Gamma(\kappa)} \int_0^\delta e^{-\mu x} x^{\kappa-1} dx.$$

Now according to our notation $\ln(1-p) = G(\mu, \kappa)$,

$$\begin{aligned} \frac{-dp}{1-p} &= \frac{\partial G}{\partial \mu} d\mu + \frac{\partial G}{\partial \kappa} d\kappa, \\ \frac{V(p)}{(1-p)^2} &= \left(\frac{\partial G}{\partial \mu}\right)^2 V(\mu) + \left(\frac{\partial G}{\partial \kappa}\right)^2 V(\kappa) + 2\left(\frac{\partial G}{\partial \mu}\right)\left(\frac{\partial G}{\partial \kappa}\right) \text{cov}(\mu, \kappa). \end{aligned}$$

For finding variance and covariances of μ and κ we have

$$\begin{aligned} \begin{pmatrix} v(\mu) & \text{cov}(\mu, \kappa) \\ \text{cov}(\mu, \kappa) & v(\kappa) \end{pmatrix} &= \begin{pmatrix} -E\left(\frac{\partial^2 L}{\partial \mu^2}\right) & -E\left(\frac{\partial^2 L}{\partial \mu \partial \kappa}\right) \\ -E\left(\frac{\partial^2 L}{\partial \mu \partial \kappa}\right) & -E\left(\frac{\partial^2 L}{\partial \kappa^2}\right) \end{pmatrix}^{-1}, \\ \begin{pmatrix} (2179.4)m & (-87.2)m \\ (-87.2)m & (14.3)m \end{pmatrix}^{-1} &= \begin{pmatrix} \frac{0.0006}{m} & \frac{-0.0037}{m} \\ \frac{-0.0037}{m} & \frac{0.0924}{m} \end{pmatrix}, \end{aligned}$$

from which we get

$$V(p) = \frac{0.477}{m} = 0.0011.$$

Thus we find $t=0.24$ which gives no evidence of a concentration of probability at zero.

The distribution of mean rainfall based on 94 years from this type III population is given by

$$f(\bar{x}) = \frac{(n\mu)^{n\kappa}}{\Gamma(n\kappa)} e^{-n\mu\bar{x}} \bar{x}^{n\kappa-1} = \frac{(1.128)^{9.870}}{\Gamma(9.870)} e^{-1.128\bar{x}} \bar{x}^{8.870}.$$

It corresponds approximately to a χ^2 distribution with 10 d.f. with skewness

$$\frac{1}{2} \frac{\mu_3}{\sigma^3} = 0.318.$$

Thus the distribution of the mean of 94 observations is far from normal and a test of significance of such a mean, based on a normal distribution, may be quite misleading.

IV. REFERENCES

- BOWEN, E. G. (1953).—*Aust. J. Phys.* **6**: 490-7.
 COHEN, A. C., JR. (1950).—*J. Amer. Statist. Ass.* **45**: 411-23.
 DES RAJ (1953).—*J. Amer. Statist. Ass.* **48**: 336-49.
 HANNAN, E. J. (1955).—*Aust. J. Phys.* **8**: 289-97.
 PEARSON, K. (1922).—"Tables of the Incomplete Γ -Function." (Cambridge Univ. Press.)

SHORT COMMUNICATIONS

NOTE ON PLASMA-ELECTRON OSCILLATIONS*

By K. G. EMELEUS† and T. K. ALLEN†

Wild, Murray, and Rowe (1954) have recently presented strong evidence that coronal plasma oscillations are responsible for certain forms of solar radio noise. It may be of interest in this connexion to mention briefly some results, partly new, which we have obtained in a laboratory study of plasma oscillations. The interpretation of the experiments is not in all respects unambiguous, and we have not attempted any quantitative scaling from discharge tube to solar conditions—nor to the apparently allied problem of colliding nebulae—but there is at least some parallelism between the phenomena probably occurring in the different sources. The experiments have been done with hot-cathode discharges in mercury vapour or argon, at pressures of order of 10^{-3} mm Hg, with cathode falls of potential between the ionization potential and a little over 100 V, i.e. with primary electrons moving with velocities rather greater than 10^8 cm/sec. The oscillations have been mainly studied with a probe inserted into the discharge and coupled to an external variable-frequency resonator. So far as the detecting system does not produce distortion, the spectrum obtained is thus comparable with Wild, Murray, and Rowe's "natural spectrum".

It has been found that :

(1) There is no strong coherent oscillation of any large volume of plasma, with frequency near the Langmuir plasma-electron frequency (f), in the tubes we have used.

(2) Oscillations with frequency close to f occur very commonly in limited regions where the plasma is traversed by one or more beams of primary electrons.

(3) The conditions under which they are produced are consistent with their maintenance through some type of slipping stream interaction, growing in amplitude along one of the beams, and rapidly becoming catastrophic through the operation of Bohm and Gross's (1949) process of electron trapping.

(4) Contrary to what was thought earlier (Emeleus and Neill 1951), harmonics can be generated, i.e. frequencies $2f$ and $3f$ picked up in addition to the fundamental frequency f . The conditions for their production are not entirely understood, but generally speaking, they have been found where there is other reason to believe that the plasma or beam oscillations are non-linear. There is some evidence, which needs confirmation, for the occasional occurrence of a subharmonic with frequency $\frac{1}{2}f$.

* Manuscript received January 10, 1955.

† Physics Department, Queen's University, Belfast, Northern Ireland.

(5) Oscillating beams not only acquire a longitudinal distribution of energy (Dittmer's (1926) "scattering"), but have a strong tendency to be deviated laterally. The resulting transverse motion could provide the basis of a mechanism for generating electromagnetic radiation with field components perpendicular to the original direction of the beam.

(6) Under certain conditions plasma ion oscillations, or relaxation oscillations, with frequency less than 1 Mc/s are produced as a secondary consequence of the higher frequency (plasma electron) oscillations. Their spectrum has the characteristics of low frequency "noise" (Martin and Woods 1952). If the low frequency oscillations react on the high frequency oscillations in the plasma, the frequency of the latter will become spread. Wild, Murray, and Rowe (1954) have found it necessary to postulate a spread round f in the coronal oscillations, which might conceivably receive a contribution from a comparable mechanism.

We have not actually been able to detect electromagnetic radiation with the Langmuir frequency, or its harmonics, coming direct from a plasma in which high frequency oscillations are taking place. If present, its energy outside the tube is much less than that of the corresponding oscillatory currents which can be withdrawn by an internal probe.

References

- BOHM, D., and GROSS, E. P. (1949).—*Phys. Rev.* **75**: 1851.
 DITTMER, A. F. (1926).—*Phys. Rev.* **28**: 507.
 EMELEUS, K. G., and NEILL, T. R. (1951).—*Proc. R. Irish Acad.* **53**: 197.
 MARTIN, H., and WOODS, H. A. (1952).—*Proc. Phys. Soc. Lond. B* **65**: 281.
 WILD, J. P., MURRAY, J. D., and ROWE, W. C. (1954).—*Aust. J. Phys.* **7**: 439.

A VARIABLE ENERGY CYCLOTRON*

By D. E. CARO,† L. H. MARTIN,† and J. L. ROUSE†

In the past, variable energy particle accelerators have, in general, been limited to the energy range below about 4 MeV. Many fixed energy machines have been built to operate above this energy but the field of physical research made possible by a variable energy machine between 2 and 12 MeV has been largely neglected. With particles of energy above about 10 MeV some energy variation may be achieved with absorbing foils, but this technique is unsatisfactory at lower energies. In consequence it was felt that the development of a machine capable of accelerating protons in the energy range 2–12 MeV would make a useful contribution to the tools of the nuclear physicist.

Electrostatic generators are costly and difficult to construct for energies much above 3 MeV, and early in 1953 it was decided to make design studies

* Manuscript received March 1, 1955.

† Physics Department, University of Melbourne.

of a variable energy linear accelerator and of a variable energy cyclotron. In the middle of the year the decision was taken to build the cyclotron. The choice was determined by financial considerations, by technical difficulties, and by the greater flexibility of the cyclotron. The variable energy cyclotron described here is designed to accelerate protons in the range 2–12.5 MeV, deuterons in the range 4–6.3 MeV, and α -particles in the range 8–12.5 MeV. The machine can be used not only as a flexible research instrument but also for the production of radioactive isotopes. The obligation to supply this latter facility for medical research and treatment in this country has strongly influenced some details of the design. The machine has some unusual technical features and it was thought that these are of sufficient interest to justify a brief description though the machine is not at present operational.

Technical Details

In order to make a variable energy cyclotron with the range specified, it is necessary to be able to vary both the dee voltage frequency and the magnetic field over a range of 2.7 to 1.

The Radio-frequency System.—The radio-frequency system is required to deliver 100 kV to the dee with continuously variable frequencies in the range 8.5–22 Mc/s. To cover this tuning range with a conventional quarter wave line, the dee stem would have a length of about 14 ft and a short circuit which could be moved over a distance of nearly 10 ft. With such a long stem supporting the dee, it was certain that difficulty would be experienced in keeping the dee central within the vacuum box. Accordingly a compromise solution has been adopted in which the dee stem is made some 8 ft long with the short circuit variable over a distance of about 5 ft. To obtain resonance at low frequencies a variable loading capacitor has been installed on the dee stem just back from the dee. This allows the effective dee capacitance to be increased fourfold for tuning to the lowest frequency.

If a small frequency change is to be made, the dee line is tuned from the control desk using the loading capacitor. For a large frequency change, the short circuit is moved. The short-circuit spider consists of four wide copper straps terminated by fingers which make contact with the inner and outer conductors. The fingers are pressed against the line conductors by shoes, actuated by eight hydraulic cylinders operating at a pressure of 200 lb/in².

The radio-frequency power (up to 50 kW) is supplied by an oscillator-power amplifier system, loop coupled to the dee line. The final power amplifier is a neutralized single triode stage using an STC type 3Q/261E valve with a plate dissipation of 20 kW.

The variable radio-frequency requirement dictates the choice of a single dee machine. The 3 in. dee aperture is comparatively large to permit the acceleration of beam currents of the order of 1 mA for isotope production. The deflector is contained within the dee and is supplied with up to 80 kV D.C. provided by a voltage doubling rectifier set. Figure 1 is a line drawing of the main features of the machine.

The Design and Performance of the Magnet.—The magnet is made from cast steel, with 0·10 per cent. of carbon present and less than 0·15 per cent. of silicon and 0·15 per cent. of manganese. It weighs 45 tons and is 9 ft 4 in. long, 4 ft wide and 7 ft 3 in. high. The pole pieces are 40 in. in diameter and the

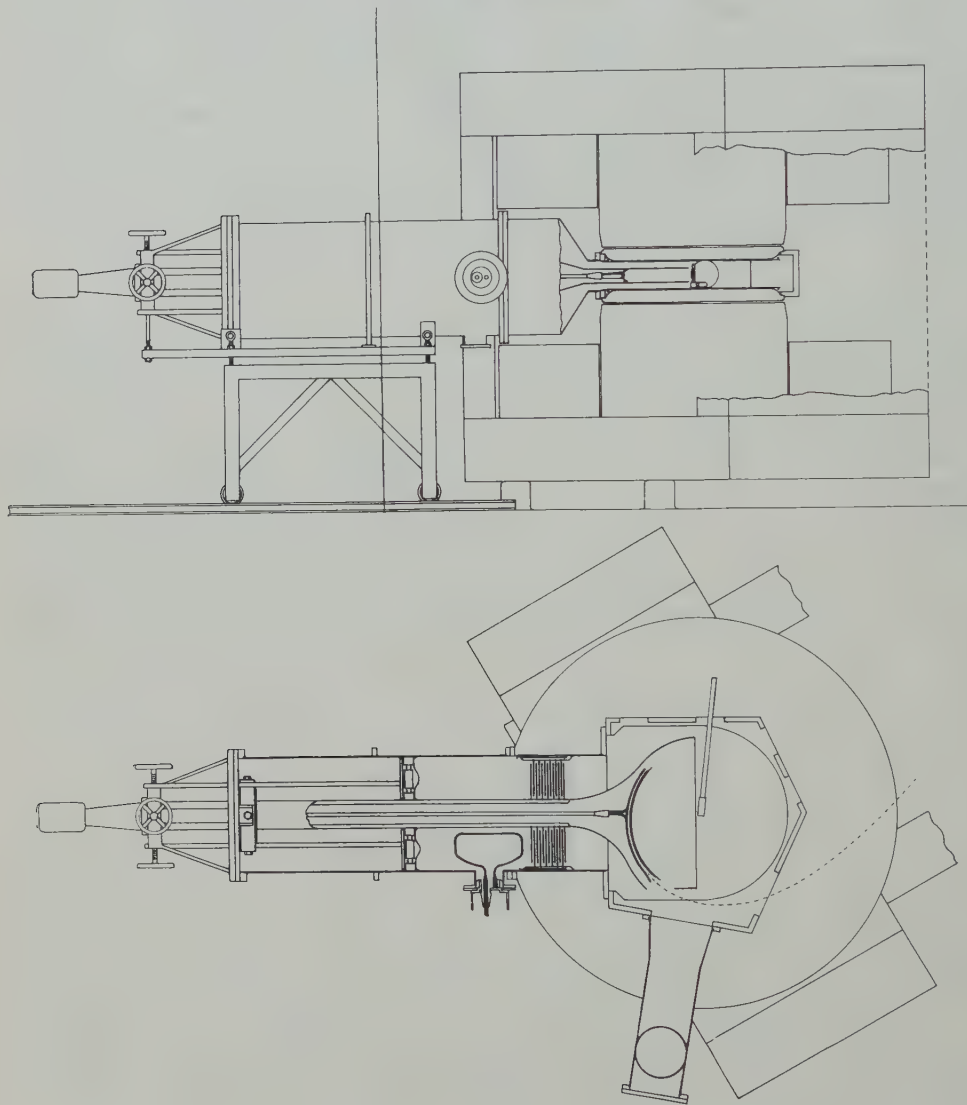


Fig. 1.—Design features of variable energy cyclotron.

magnet gap is 6 in. wide. The pole tips are 2·7 in. thick and form the lids of the vacuum box. Between the pole tips and the main pole pieces there is a $\frac{5}{16}$ in. shim gap in which iron shims may be placed to give fine adjustment to the field in the main gap. The pole tips are held in place with 2 in. steel bolts through the pole pieces.

The design of the pole tips differs from that of other cyclotrons where the final energy of the accelerated particles is held constant. In order to vary the particle energy it is necessary to change the magnetic field in the gap while keeping the field profile constant. This requires that the iron in the pole tips be kept below saturation particularly along the gap faces. For this reason the pole tips were designed with a rounded outer edge in such a way that the maximum flux density at any point of the surface of the iron was less than 1.3 times the flux density in the gap, thus avoiding the saturation that occurs at sharp edges. The design field profile called for a decrease in the field from the centre to the exit radius at 14.4 in. of 1.8 per cent. The maximum desired field was 14,000 G.

The machining tolerances on the magnet are critical. The two pole tips need to be accurately machined and the final gap between the pole tips was specified to ± 0.002 in. In fact with the pole tips in position the gap was found to be accurate to ± 0.0015 in. Small tolerances are required because the field should have circular symmetry about the centre of the gap with errors of the order of 0.01 per cent.

The field at the centre of the gap was measured as a function of excitation current using a small coil and flux meter calibrated against a nuclear resonance signal. A field of 12,500 G was obtained at 800 A at which point the excitation curve showed little signs of saturation. A current of 1200 A gave a field of 14,500 G with considerable saturation present. The field profile has been investigated using a nuclear resonance system and a pair of rotating coils. These measurements showed that the profile closely followed the design, that the profile was independent of the central field for fields up to 12,500 G, and that there was little change in the profile for fields up to 14,000 G. The pair of rotating coils was used to measure the symmetry of the field about the central axis at various radii. The results obtained showed the azimuthal variations out to the exit radius to be less than ± 0.02 per cent.

The coils for exciting the magnet are wound with 256 turns of hollow rectangular aluminium conductor. These coils are energized by a motor generator which is capable of providing 1300 A at 70 V. The field is stabilized to 1 part in 5000 by electronic stabilization of the current. The coils are cooled by passing 8 gal of water per minute through the conductor.

We are grateful for financial assistance from a large number of outside organizations, notable among which are: The Anti-Cancer Council of Victoria; Australian Paper Manufacturers Ltd.; Thomas Baker (Kodak), Alice Baker, and Eleanor Shaw benefactions; The Chamber of Manufactures, Melbourne; Ducon Condenser Co.; The Herald and Sun Newspapers; Standard Telephones and Cables Pty. Ltd.; The Steel Company of Australia Ltd.; Thomsons (Castlemaine) Ltd.

The assistance of W. G. Caldwell, G. J. F. Legge, J. R. Moroney, A. M. Segar, H. H. Thies, and H. A. Waters in the construction of the machine is gratefully acknowledged.

EXCITATION STUDIES OF REACTIONS OCCURRING IN THE PROTON BOMBARDMENT OF ^{10}B †

By J. W. G. WIGNALL‡

This note reports an investigation of the excitation functions for the reactions $^{10}\text{B}(p,\gamma)^{11}\text{C}$ and $^{10}\text{B}(p,\alpha)^7\text{Be}$ for proton energies of up to 775 keV. Curran, Dee, and Petržílka (1939) investigated the (p,γ) reaction by analysing the decay curve of activations produced in the bombardment of a thick boron target by 960 keV protons; they found no evidence for this reaction and quoted an upper limit of 10^{-13} ^{11}C nuclei formed per incident proton. Later, Walker (1950), using a magnetic pair spectrometer, observed a (9.47 ± 0.12) MeV γ -ray emitted in the bombardment of a thick separated ^{10}B target by 1200 keV protons, which was attributed to the reaction $^{10}\text{B}(p,\gamma)^{11}\text{C}$. The yield was 3.4×10^{-10} γ -rays per proton.

The excitation function for the 9.5 MeV γ -radiation has previously been measured, using NaI(Tl) single crystal spectrometers, by three groups. Krone and Seagondollar (1953) studied the proton energy region 0.6–1.4 MeV, and found indications of a resonance at about 780 keV. Day and Huus (1954), studying the energy region 0.5–2.4 MeV, failed to observe this resonance, but found one at 1.2 MeV. Recent work by Hahn, Kern, and Farney (1955) supports the result of Day and Huus.

All of these (p,γ) excitation curve experiments suffer from a large background of 4, 12, and 16 MeV γ -rays produced from the reaction $^{11}\text{B}(p,\gamma)^{12}\text{C}$ due to residual ^{11}B impurity present even in the separated targets used. Because of this, it is difficult both to observe the spectrum of γ -rays from $^{10}\text{B}(p,\gamma)^{11}\text{C}$ and to separate its contribution to the γ -ray spectrum in an excitation curve measurement.

The method used in the present experiment was to measure the yield of ^{11}C nuclei formed in proton bombardment of a thick boron-containing target, as a function of proton energy. The ^{11}C nuclei were detected by means of their decay to ^{11}B with the emission of 0.98 MeV positrons: half-life 20.4 min (Ajzenberg and Lauritsen 1952). The presence of ^{11}B in the target is then of no consequence, and targets containing natural boron were used, their only disadvantage being the smaller percentage of ^{10}B which they contain. The use of such an activation method will in principle give all resonances in which ^{11}C is formed, irrespective of the type of γ -ray transition preceding the formation of ^{11}C in its ground state.

The (p,α) excitation curve was studied simultaneously with the (p,γ) curve by measurement of the yield of ^7Be as a function of proton energy, the ^7Be

† Manuscript received February 21, 1955.

‡ Physics Department, University of Melbourne.

(5.3 day half-life) being detected by means of the 0.478 MeV γ -radiation emitted in the electron capture decay to ${}^7\text{Li}$ (Ajzenberg and Lauritsen 1952). The (p, α) curve has been previously studied, through observation of the α -particles emitted, by Burcham and Freeman (1949, 1950) who found no resonances in the region 200–800 keV. The (p, α) activation experiment reported here is of a confirmatory nature.

In each run of the present experiment a thick target of borax, prepared by melting about 0.7 g of anhydrous borax ($\text{Na}_2\text{B}_4\text{O}_7$) into a $\frac{3}{8}$ in. diameter nickel target cup, was bombarded for 40 min (about two half-lives of ${}^{11}\text{C}$) by an analysed beam of protons from the 750 keV electrostatic generator at the University of Melbourne. The current received by the target passed first through an "ordinary" integrator to record the total charge received and then into a "leaky" integrator, of the type used by Snowdon (1950), having a decay time constant made equal to that for the ${}^{11}\text{C}$ decay. This recorded the effective charge for ${}^{11}\text{C}$ production allowing for the fact that the ${}^{11}\text{C}$ was decaying with a half-life comparable to the time of bombardment.

After bombardment the cup was extracted from the target holder and placed in fixed geometry in the activation counting apparatus. The relative amount of ${}^{11}\text{C}$ was measured by coincident detection of the two 0.51 MeV quanta produced by annihilation of the positrons in the surrounding material, using a pair of NaI(Tl) single crystal γ -ray spectrometers feeding into a coincidence circuit of 1.0 μsec resolving time. The quantity of ${}^{11}\text{C}$ was determined by analysis of the coincidence decay curve, taken over a period of about three half-lives after bombardment; division by the leaky integrator reading then gave the relative yield of ${}^{11}\text{C}$ (the background to be subtracted varied from run to run owing to a small contribution of coincidences arising from "double scattering" of 0.478 MeV γ -rays from ${}^7\text{Be}$). The relative yield of ${}^7\text{Be}$ was given by the amount of long-lived γ -ray activation per proton, determined at a time when the ${}^{11}\text{C}$ had decayed to a negligible amount.

In addition to the usual procedures employed in order to obtain a quantitatively reliable excitation curve (suitable beam/target geometry, secondary electron suppression, stability of activation counting apparatus), the following special precautions had to be taken to keep the conditions of bombardment and counting constant from run to run. (i) It was found that the beam slowly bored a hole through the borax target material; this effect was minimized by the use of low beam currents (10–15 μA), and a sufficiently large thickness of borax was used. During each bombardment the target condition was checked every few minutes by means of a NaI(Tl) γ -ray monitor, with bias set at 3 MeV. This would detect any reduction of target thickness below that required to stop the incident protons completely; such an effect would appear as a drop in the prompt γ -ray yield from ${}^{11}\text{B}(p, \gamma){}^{12}\text{C}$ and to a smaller extent from ${}^{23}\text{Na}(p, \gamma){}^{24}\text{Mg}$ and ${}^{10}\text{B}(p, \gamma){}^{11}\text{C}$. (ii) It was observed that some activated material was being evaporated, and so lost from the target, by the heating due to the beam and that this material condensed on the sides of the target cup; in order to correct this a lining of aluminium foil was placed around the sides of the cup during bombardment so as to collect this material; after bombardment the foil was crushed back

on to the target, thus retaining all activation in a small volume for counting purposes.

Precautions were taken to eliminate any interference by unwanted activations due to primary proton reactions and secondary photodisintegration or slow neutron reactions with target material, cup, or counters. The most serious of these activations was ^{13}N , a positron emitter with 10 min half-life (Ajzenberg and Lauritsen 1952) produced by the reaction $^{12}\text{C}(p,\gamma)^{13}\text{N}$ in carbon deposited on the target from oil vapours in the system; the use of a liquid oxygen trap, and the heating of the target to about 200 °C, reduced this to negligible amount.* Another activation to be guarded against was ^{128}I , which, it was observed, could be produced by the capture of slow neutrons by the iodine in the detecting crystals; this could give β - γ coincidences with a 25 min half-life. The activation counting apparatus was therefore placed at some distance from the target: experiments without the target in place in this apparatus, and the observed half-life of the activation (see Fig. 1), proved conclusively that the short-lived activation being detected was not due to this process.

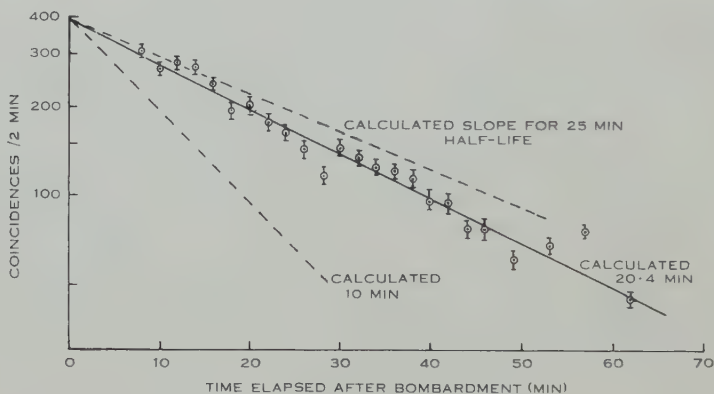


Fig. 1.—Typical coincidence decay curve obtained after proton bombardment of borax target (proton energy 750 keV), background subtracted.

A typical coincidence decay curve obtained in the experiment is shown in Figure 1, on which the calculated slopes for 10, 20.4, and 25 min half-lives have been drawn. Although the statistical fluctuations are large, it is clear that the activity measured arises from the decay of ^{11}C . A decay curve, taken over a period of weeks, for the intensity of γ -radiation from a bombarded target exhibited a half-life of 53 days as expected from ^7Be .

The excitation curves obtained are shown in Figure 2. Figure 2 (a) represents the prompt γ -ray yield mentioned above, and acts as a useful check on the target condition and the reliability of the current monitoring apparatus.

* At an early stage amorphous boron was tried as a target material but was found to contain a large proportion of carbon. There was also evidence that the effective percentage of ^{10}B in the surface layers of a target pressed from this material was changed by bombardment.

The excitation curve for $^{10}\text{B}(p, \alpha)^7\text{Be}$ is given in Figure 2 (b); within the limitations imposed by the fact that the ^7Be count is only of order one-tenth background this exhibits the smooth rise expected from the α -particle measurements of Burcham and Freeman. Figure 2 (c) is the thick target excitation function for $^{10}\text{B}(p, \gamma)^{11}\text{C}$; this is seen to have a smooth rise up to about 750 keV, with no observed resonances within the statistical accuracy of the measurements.

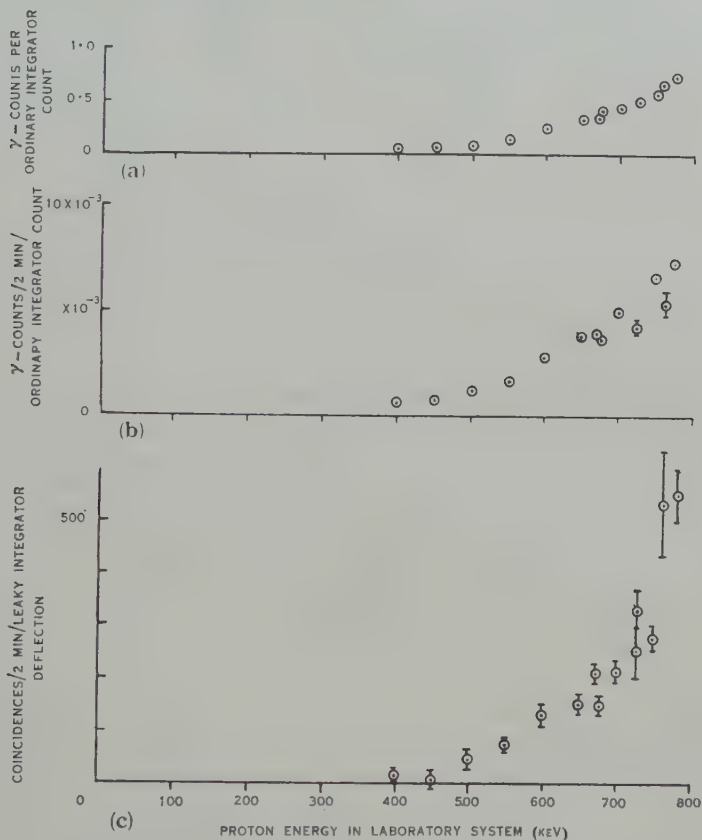


Fig. 2.—Thick target excitation functions for reactions of protons with borax target. (a) Prompt γ -rays > 3 MeV; (b) $^{10}\text{B}(p, \alpha)^7\text{Be}$; (c) $^{10}\text{B}(p, \gamma)^{11}\text{C}$.

No attempt has been made at accurate measurement of absolute yields of ^{11}C and ^7Be , but rough estimates of counting efficiencies lead to thick borax target yields, at 700 keV, of 5×10^{-12} ^{11}C nuclei per proton and 3×10^{-8} ^7Be nuclei per proton.

The author is indebted to Professor L. H. Martin for advice and encouragement throughout the course of this work, and to Dr. J. R. Prescott, Mr. J. G. Campbell, and Dr. D. N. F. Dunbar for guidance in some of the experimental problems.

References

- AJZENBERG, F., and LAURITSEN, T. (1952).—*Rev. Mod. Phys.* **24**: 321.
 BURCHAM, W. E., and FREEMAN, JOAN M. (1949).—*Phil. Mag.* **40**: 807.
 BURCHAM, W. E., and FREEMAN, JOAN M. (1950).—*Phil. Mag.* **41**: 337.
 CURRAN, S. C., DEE, P. I., and PETRŽÍLKA, V. (1939).—*Proc. Roy. Soc. A* **169**: 269.
 DAY, R. B., and HUUS, T. (1954).—*Phys. Rev.* **95**: 1003.
 HAHN, T. M., JR., KERN, B. D., and FARNEY, G. K. (1955).—*Bull. Amer. Phys. Soc.* **30** (1): 44.
 KRONE, R. W., and SEAGONDOLLAR, L. W. (1953).—*Phys. Rev.* **92**: 935.
 SNOWDON, S. C. (1950).—*Phys. Rev.* **78**: 299.
 WALKER, R. L. (1950).—*Phys. Rev.* **79**: 172.

NEUTRONS EMITTED IN THE DEUTERON BOMBARDMENT OF CARBON*

By J. R. BIRD†

Energy levels in ^{14}N have been observed at 2.31, 3.95, 4.91, and 5.10 MeV in the study of the inelastic scattering of protons and deuterons by nitrogen (Bockelman *et al.* 1953), and these levels may be used to explain all the low energy γ -rays associated with the ^{14}N nucleus except for a 0.725 MeV γ -ray (Woodbury, Day, and Tollestrup 1953; Ajzenberg and Lauritsen 1955). Experiments on the reaction $^{13}\text{C}(d,n)^{14}\text{N}$ give evidence for the same levels with additional ones suggested at 3.47 MeV (Mandeville and Swann 1950) and 3.8 MeV (Benenson 1953) although the energy values obtained do not agree very well with those of Bockelman *et al.* It may be noted that the measurements of Benenson give better agreement if, in the calculation of the level energies, his measured ground state Q -value is used rather than the value given by the atomic masses.

A further study of the reaction $^{13}\text{C}(d,n)^{14}\text{N}$ has been made by bombarding a thick target of natural carbon with 0.92 MeV deuterons for an integrated target exposure of 100,000 μC . Although the more abundant isotope of carbon does not interfere with the study of energy levels in ^{14}N below 5 MeV the low yield of neutrons from the required reaction necessitates precautions to reduce the number of neutrons from contaminant reactions. For this reason the copper diaphragm used to limit the size of the deuteron beam and the copper target backing were kept at several hundred degrees C during the exposure.

The camera was constructed from 0.030 in. brass and the supports for the target mounting and the electrical leads were mounted in a side tube at a considerable distance from the target in order to reduce the effects of neutron scattering. A paraffin wax stack was mounted between the camera and the electrostatic generator to cut down the number of neutrons from deuterium

* Manuscript received March 4, 1955.

† Physics Department, University of Melbourne.

contamination of the accelerating tube and the back scattering of neutrons by the analysing magnet.

Pieces, 1 by 3 in., of 400 μ Ilford C2 plates were held by wire frames at angles of 0, 30, 90, and 150° to the direction of the incident deuteron beam. The plane of the surface of each emulsion lay in the direction of the neutron flux and the front edges of the plates were at a distance of 3 cm from the target. The plate chamber was isolated from the target chamber and filled with air which was kept dry with phosphorus pentoxide. The plates were vacuum dried for 24 hr before loading in the camera, and measurements of the emulsion density for plates which were treated in a similar way gave the value 4.00 g/cm³. The plates were processed by the "temperature" method in order to ensure uniform development throughout the depth of the emulsion.

Measurements were made on each plate over areas subtending an angle of $\pm 2^\circ$ at the target, and tracks were accepted which had an angle of dip in the unprocessed emulsion not greater than $\pm 5^\circ$ and an angle in the plane of the emulsion not greater than $\pm 10^\circ$. These angles together with the distance between the first and last grains were tabulated for each acceptable track which did not cross either surface of the emulsion. The measured range of each recoil proton was corrected according to the angle of recoil and then converted to neutron energy using the microscope calibration and a range-energy relation based on the results of Rotblat (1951). The energy distributions plotted in 100 keV intervals are shown in Figure 1, after correction for the variation of the neutron-proton scattering cross section with neutron energy. No additional correction has been made for the loss of tracks from the emulsion since this would change the correction factor by less than 2 per cent. A total of about 3000 tracks was measured and the ordinates in Figure 1 represent the true variation of neutron intensity with energy and angle of emission.

Energy Levels

The peak energies obtained from Figure 1 have been corrected to give the corresponding thin target energies, making allowance for the possible effects of resonances in the yield of neutrons as observed in the reaction $^{13}\text{C}(d,n)^{14}\text{N}$ by Richardson (1950). The dotted peaks in Figure 1 are attributed to the D-D reaction, and give a weighted mean Q -value for this reaction of 3.24 ± 0.05 MeV, which is in satisfactory agreement with the value calculated from the atomic masses. The low energy group in the 0° spectrum is attributed to the reaction $^{12}\text{C}(d,n)^{13}\text{N}$. The number of tracks is very much underestimated in Figure 1, since they were only measured over a small portion of the area scanned. The high intensity and the estimated Q -value (-0.33 ± 0.06 MeV) are consistent with the allocation of this group to the ^{12}C reaction.

The peaks numbered 1-4 arise from the ^{13}C reaction and give weighted mean Q -values of 5.325 ± 0.04 , 3.02 ± 0.05 , 1.37 ± 0.03 , 0.37 ± 0.02 MeV, which define levels in ^{14}N at 2.30, 3.95, and 4.95 MeV. The probable errors represent the observed fluctuations in individual Q -values together with the uncertainties in the range-energy relation and bombarding energy. Peak 4 is about 100 keV wider than would be expected from the width of the higher

energy peaks, and this can be explained by the assumption of a pair of levels at 5 MeV excitation. The 30 and 90° peaks are wider than the 0° peak which may support the suggestion by Benenson (1953) that the angular distributions of the two groups are different.

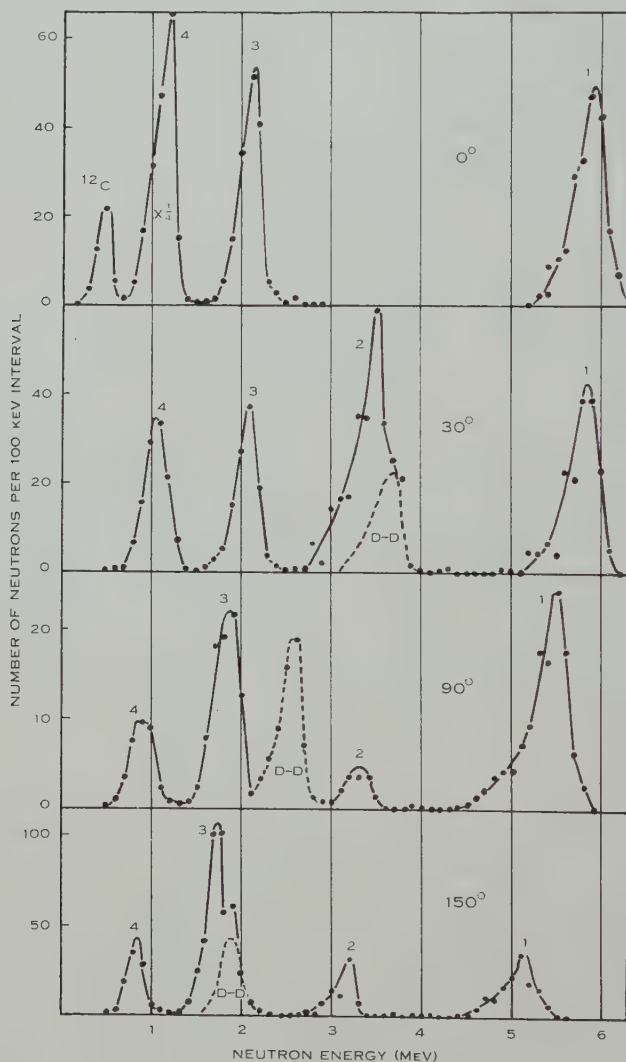


Fig. 1.—Neutron spectra from the reaction $^{13}\text{C}(d,n)^{14}\text{N}$.

These results are in satisfactory agreement with those obtained from other reactions leading to the same final nucleus. No evidence is obtained for the existence of a level at 3.47 MeV as reported by Mandeville and Swann (1950), and an upper limit of 5 per cent. of the ground state intensity can be placed on the possible intensity of a neutron group corresponding to such a level. Similarly no indication is obtained of a doublet at 4 MeV excitation as suggested by

Benenson. The lack of intermediate low-lying levels requires the allocation of the 0.725 MeV γ -ray to a cascade involving a level above 5 MeV excitation as suggested by Woodbury, Day, and Tollestrup (1953).

Angular Distributions.

The contribution to the yield of deuteron induced reactions by the process of deuteron stripping is dependent on the deuteron energy and therefore cannot be readily calculated for a thick target. However, the forward peaks which are characteristic of the stripping process will still be observed and may give qualitative information about the transfer of angular momentum to the final nucleus. The variation with angle in the laboratory system of coordinates, of

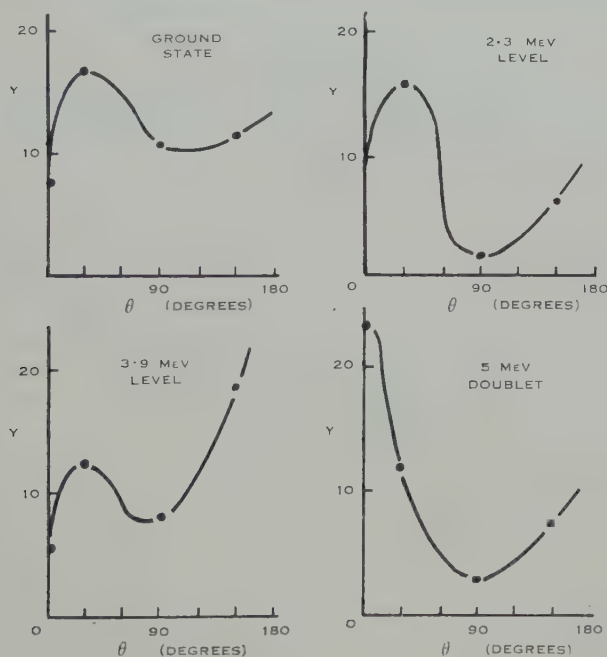


Fig. 2.—Variation with angle of observation of the thick target yield of neutrons from natural carbon. Y , number of neutrons per unit solid angle per 10^{10} deuterons; θ , angle in laboratory system of coordinates.

the thick target yield of the neutron groups corresponding to each state of ^{14}N , is shown in Figure 2. Each curve has a pronounced forward peak which suggests that stripping makes an appreciable contribution to the yield of neutrons for deuteron energies below 0.92 MeV.

The angular distributions observed for the excited states are similar to those given by Benenson (1953) and would therefore lead to the same conclusions as to the spin and parity of these states. However, the distribution for the ground state has a peak at an angle greater than 0° and agrees with the results of Bromley and Goldman (1952) rather than those of Benenson. This implies

even parity for the ground state of ^{14}N which is therefore different from the ground state of ^{14}C and leaves unexplained the forbidden nature of the β -decay of ^{14}C .

Thanks are due to Professor L. H. Martin and Dr. J. C. Bower for their interest in this work, and to the various members of the electrostatic generator team for assistance in making the exposures.

References

- AJZENBERG, F., and LAURITSEN, T. (1955).—*Rev. Mod. Phys.* **27** : 77.
BENENSON, R. E. (1953).—*Phys. Rev.* **90** : 420.
BOCKELMÁN, C. K., BROWNE, C. P., BUECHNER, W. W., and SPERDUTO, A. (1953).—*Phys. Rev.* **92** : 665.
BROMLEY, D. A., and GOLDMAN, L. M. (1952).—*Phys. Rev.* **86** : 790.
MANDEVILLE, C. E., and SWANN, C. P. (1950).—*Phys. Rev.* **79** : 787.
RICHARDSON, J. E. (1950).—*Phys. Rev.* **80** : 850.
ROTLAT, J. (1951).—*Nature* **167** : 550.
WOODBURY, H. H., DAY, R. B., and TOLLESTRUP, A. V. (1953).—*Phys. Rev.* **92** : 1199.

CORRIGENDUM

VOLUME 7, NUMBER 3

Page 410, line 7 of Summary : *For* $<5 \times 10^{-8}$ sec *read* $>5 \times 10^{-8}$ sec

HYDROMAGNETIC STABILITY OF A CURRENT LAYER

By R. E. LOUGHHEAD*

[Manuscript received January 19, 1955]

Summary

The hydromagnetic stability of a uniform current flowing along a magnetic field and confined within a pair of parallel planes is discussed by the method of normal modes. The condition for marginal stability is derived and discussed with reference to two special cases.

It is also shown that the velocity of Alfvén waves along a magnetic field in a region bounded by parallel planes is reduced due to the inertia of the surrounding medium.

I. INTRODUCTION

In a recent paper Dungey and Loughhead (1954) have considered the stability of a current of uniform density flowing coaxially along a magnetic field contained within an infinitely long cylinder, the resultant magnetic field having twisted lines of force. The present paper deals with the similar problem of the stability of a uniform current flowing in a perfectly conducting fluid within a region bounded by parallel planes. Using the principle of exchange of stabilities the condition for marginal instability is derived, and then discussed in some detail for two special cases.

The first is that of the stability of a uniform current flowing along a magnetic field, the current and field both vanishing outside a region bounded by a pair of parallel planes. It is concluded that, for any layer of finite thickness, there are modes of disturbance for which the current distribution is unstable.

The second case is an idealization of the "pinch" effect, in which the magnetic field in the fluid is everywhere perpendicular to the current. It is found that the system possesses unstable modes of disturbance when the magnitude of the field inside exceeds that of the field outside the current layer. However, in conformity with the result of Kruskal and Schwarzschild (1954), the current system is shown to be stable so long as the magnitude of the field outside exceeds that inside the layer.

II. EQUATIONS FOR SMALL PERTURBATIONS

Suppose that, referred to Cartesian axes Ox, y, z , a uniform current flows along an imposed magnetic field H_z in the region bounded by the planes $x = \pm a$. Associated with the current there is a transverse component of the magnetic field

$$H_y = A_x, \quad \dots\dots\dots (1)$$

where A is a constant related to the current density. Outside the current layer the magnetic field is taken to have the uniform transverse components

$$\left. \begin{array}{ll} H_y = A_+, & \text{for } x > a, \\ A_-, & \text{for } x < -a. \end{array} \right\} \dots\dots\dots (2)$$

* Division of Physics, C.S.I.R.O., University Grounds, Sydney.

For generality the components of the magnetic field in the direction of the current are written

$$\left. \begin{aligned} H_z &= B, & \text{for } -a < x < a, \\ &= B_+, & \text{for } x > a, \\ &= B_-, & \text{for } x < -a, \end{aligned} \right\} \dots\dots\dots (3)$$

where B , B_+ , and B_- are all constants.

The results of small perturbations of the current system are therefore described by the hydromagnetic equations for an ionized fluid in the presence of a steady magnetic field of the form

$$\mathbf{H} = [0, H_y(x), H_z(x)], \dots\dots\dots (4)$$

where $H_y(x)$ and $H_z(x)$ are specified functions of x . Since it may be shown that the stability does not depend on either the density or compressibility of the fluid (see Appendix I) it is supposed that the medium is incompressible and has the uniform density μ . The electrical conductivity of the fluid is taken to be infinite.

In Gaussian units the equations governing small perturbations \mathbf{h} and \mathbf{v} in the magnetic field and fluid velocity, respectively, may be written in the forms :

$$\frac{\partial \mathbf{h}}{\partial t} = (\mathbf{H} \cdot \text{grad}) \mathbf{v} - (\mathbf{v} \cdot \text{grad}) \mathbf{H}, \dots\dots\dots (5)$$

$$4\pi\mu \frac{\partial \mathbf{v}}{\partial t} = -\text{grad } \psi + (\mathbf{H} \cdot \text{grad}) \mathbf{h} + (\mathbf{h} \cdot \text{grad}) \mathbf{H}, \dots\dots (6)$$

$$\text{div } \mathbf{v} = 0, \dots\dots\dots (7)$$

where $\psi/4\pi$ denotes the variation in the total pressure (gas pressure plus $\mathbf{H}^2/8\pi$).

Applying the method of normal modes solutions of these equations are sought in which the variables are proportional to

$$\exp i(\omega t + my + nz), \dots\dots\dots (8)$$

where ω , m , and n refer to an arbitrary mode. Inserting (8) in (5), (6), and (7) and defining

$$K = mH_y + nH_z, \dots\dots\dots (9)$$

the following Cartesian equations are obtained :

$$i\omega h_x = iKv_x, \dots\dots\dots (10)$$

$$i\omega h_y = iKv_y - \frac{\partial H_y}{\partial x} \cdot v_x \dots\dots\dots (11)$$

$$i\omega h_z = iKv_z - \frac{\partial H_z}{\partial x} \cdot v_x, \dots\dots\dots (12)$$

$$4\pi\mu i\omega v_x = -\frac{\partial \psi}{\partial x} + iKh_x, \dots\dots\dots (13)$$

$$4\pi\mu i\omega v_y = -im\psi + iKh_y + \frac{\partial H_y}{\partial x} \cdot h_x, \dots\dots\dots (14)$$

$$4\pi\mu i\omega v_z = -in\psi + iKh_z + \frac{\partial H_z}{\partial x} \cdot h_x, \dots\dots\dots (15)$$

$$\frac{\partial v_x}{\partial x} + imv_y + inv_z = 0. \dots\dots\dots (16)$$

After some reduction one obtains the first order equations

$$(4\pi\mu\omega^2 - K^2)\frac{\partial v_x}{\partial x} - i\omega(m^2 + n^2)\psi = 0, \quad \dots\dots\dots (17)$$

$$\omega\frac{\partial\psi}{\partial x} + i(4\pi\mu\omega^2 - K^2)v_x = 0, \quad \dots\dots\dots (18)$$

in the two variables v_x and ψ . Clearly, from equations (17) and (18), v_x and ψ must be continuous across the boundary planes $x = \pm a$.

By elimination it is found that inside the region $-a < x < a$, where K takes the value

$$K = mA_x + nB, \quad \dots\dots\dots (19)$$

ψ is determined by the equation

$$\frac{\partial^2\psi}{\partial x^2} + \frac{2mA_x}{4\pi\mu\omega^2 - K^2} \cdot \frac{\partial\psi}{\partial x} - (m^2 + n^2)\psi = 0. \quad \dots\dots\dots (20)$$

Outside this region, where K assumes the values

$$\left. \begin{aligned} K_+ &= mA_+ + nB_+, \\ K_- &= mA_- + nB_-, \end{aligned} \right\} \quad \dots\dots\dots (21)$$

according as $x > a$ or $x < -a$, ψ is governed by the equation

$$\frac{\partial^2\psi}{\partial x^2} - (m^2 + n^2)\psi = 0, \quad \dots\dots\dots (22)$$

Before proceeding to the critical case $\omega = 0$, the above analysis may be applied to a discussion of the propagation of Alfvén waves along a uniform magnetic field H_z contained wholly within the region bounded by the planes $x = \pm a$.

III. PROPAGATION OF ALFVÉN WAVES IN A REGION BOUNDED BY PARALLEL PLANES

In this section the discussion is generalized to allow for different densities μ_1 inside, and μ_2 outside, the region bounded by parallel planes. To describe the propagation of Alfvén waves in the region between the planes $x = \pm a$ pervaded by the uniform magnetic field H_z let the current in the z -direction be zero. Thus $A = 0$, and equation (20) becomes

$$\frac{\partial^2\psi}{\partial x^2} - \kappa^2\psi = 0, \quad \dots\dots\dots (23)$$

where

$$\kappa^2 = m^2 + n^2. \quad \dots\dots\dots (24)$$

Imposing the condition that ψ remain finite as $x \rightarrow \pm \infty$ the solution of (23) can be written in the form

$$\left. \begin{aligned} \psi &= C_1 e^{\kappa x} + C_2 e^{-\kappa x}, & -a < x < a, \\ &= C_3 e^{-\kappa x}, & x > a, \\ &= C_4 e^{\kappa x}, & x < -a. \end{aligned} \right\} \quad \dots\dots\dots (25)$$

where C_1, C_2, C_3 , and C_4 are arbitrary constants. With the aid of equation (18) the corresponding solution for v_x is found to be given by

$$v_x = \left. \begin{aligned} & \frac{i\omega}{4\pi\mu_1\omega^2 - n^2B^2} \kappa (C_1 e^{\kappa x} - C_2 e^{-\kappa x}), & -a < x < a, \\ & = \frac{-i\omega}{4\pi\mu_2\omega^2} \kappa C_3 e^{-\kappa x}, & x > a, \\ & = \frac{i\omega}{4\pi\mu_2\omega^2} \kappa C_4 e^{\kappa x}, & x < -a. \end{aligned} \right\} \dots (26)$$

The requirement that v_x and ψ be continuous across the planes $x=a$ and $x=-a$ imposes the following relations on the four constants C_1, C_2, C_3 , and C_4 :

$$\left. \begin{aligned} C_1 e^{\kappa a} + C_2 e^{-\kappa a} - C_3 e^{-\kappa a} &= 0, \\ C_1 e^{\kappa a} - C_2 e^{-\kappa a} + C_3 e^{-\kappa a} Q &= 0, \\ C_1 e^{-\kappa a} + C_2 e^{\kappa a} - C_4 e^{-\kappa a} &= 0, \\ C_1 e^{-\kappa a} - C_2 e^{\kappa a} - C_4 e^{-\kappa a} Q &= 0, \end{aligned} \right\} \dots (27)$$

where Q has the value

$$Q = \frac{\mu_1}{\mu_2} \left\{ 1 - \frac{V_a^2}{V^2} \right\}, \dots (28)$$

and

$$V_a = \frac{B}{\sqrt{4\pi\mu_1}} \dots (29)$$

is Alfvén's velocity, V being the velocity of propagation of waves in the z -direction. From (27), the condition for the existence of non-trivial solutions may be expressed in the form

$$\begin{vmatrix} e^{\kappa a} & e^{-\kappa a} & -e^{-\kappa a} & 0 \\ e^{\kappa a} & -e^{-\kappa a} & e^{-\kappa a} Q & 0 \\ e^{-\kappa a} & e^{\kappa a} & 0 & -e^{-\kappa a} \\ e^{-\kappa a} & -e^{\kappa a} & 0 & -e^{-\kappa a} Q \end{vmatrix} = 0. \dots (30)$$

This may be rewritten

$$Q^2(e^{2\kappa a} - e^{-2\kappa a}) + 2Q(e^{2\kappa a} + e^{-2\kappa a}) + (e^{2\kappa a} - e^{-2\kappa a}) = 0.$$

Corresponding to the two roots Q of this equation we obtain the expressions

$$V = \pm \left(1 + \frac{\mu_2}{\mu_1} \tanh \kappa a \right)^{-\frac{1}{2}} \cdot V_a, \dots (31)$$

and

$$V = \pm \left(1 + \frac{\mu_2}{\mu_1} \coth \kappa a \right)^{-\frac{1}{2}} \cdot V_a. \dots (32)$$

These relations express the reduction of the wave velocity due to the inertia of the surrounding medium. Thus equation (31) shows that the magnitude of V decreases steadily from the value V_a when $\kappa a = 0$ to the value $(1 + \mu_2/\mu_1)^{-\frac{1}{2}} V_a$ as $\kappa a \rightarrow \infty$, while equation (32) implies that the magnitude of V increases steadily from the value 0 when $\kappa a = 0$ to the limiting value $(1 + \mu_2/\mu_1)^{-\frac{1}{2}} V_a$ as $\kappa a \rightarrow \infty$. This behaviour is illustrated in Figure 1 for the special case $\mu_1 = \mu_2$.

The above discussion shows that, when the effect of boundary conditions is taken into account, the velocity of propagation of hydromagnetic disturbances is not the simple Alfvén velocity V_a . The actual velocity of any mode of disturbance is reduced, due to the inertia of the surrounding medium, by a factor which, for magnetic regions of large extent (a large), approximates to $(1 + \mu_2/\mu_1)^{-\frac{1}{2}}$.

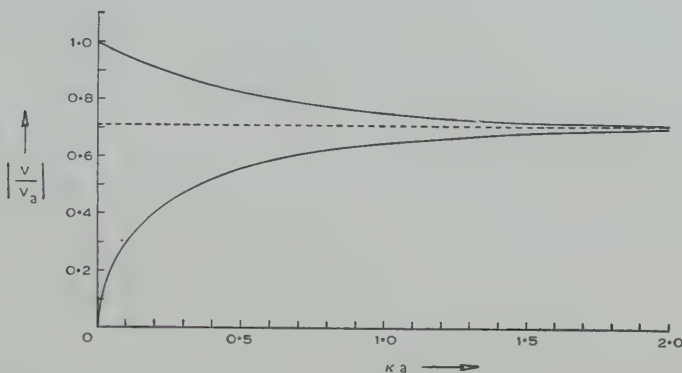


Fig. 1.—Variation of wave velocity with a for the case $\mu_1 = \mu_2$.

IV. THE CRITICAL CASE $\omega = 0$

The condition for instability in any conservative system is the existence of a solution for which the period is infinite. To determine whether any such solutions exist for different physical conditions of the system let $\omega = 0$ in the equations for the normal modes. In this case, however, equation (5) requires $\mathbf{v} = 0$, so that equation (6) must now be supplemented by the relation

$$\text{div } \mathbf{h} = 0 \quad \dots\dots\dots (33)$$

to obtain a determinate set of equations. One thus obtains the equations

$$\frac{\partial \psi}{\partial x} = iK h_x, \quad \dots\dots\dots (34)$$

$$im\psi = iK h_y + \frac{\partial H_y}{\partial x} \cdot h_x, \quad \dots\dots\dots (35)$$

$$in\psi = iK h_z + \frac{\partial H_z}{\partial x} \cdot h_x, \quad \dots\dots\dots (36)$$

$$\frac{\partial h_x}{\partial x} + imh_y + in h_z = 0. \quad \dots\dots\dots (37)$$

After some reduction these yield the first order equations

$$K \frac{\partial h_x}{\partial x} + i(m^2 + n^2)\psi = \frac{\partial K}{\partial x} \cdot h_x, \quad \dots\dots\dots (38)$$

$$\frac{\partial \psi}{\partial x} = iK h_x, \quad \dots\dots\dots (39)$$

in the two variables h_x and ψ . Equations (38) and (39) imply respectively the continuity of h_x/K and ψ across the boundary planes $x = \pm a$.

Inside the current layer where K takes the value (19), ψ is found from (38) and (39) to be governed by the equation

$$\frac{\partial^2 \psi}{\partial x^2} - \frac{2mA}{K} \cdot \frac{\partial \psi}{\partial x} - (m^2 + n^2)\psi = 0. \quad (40)$$

On changing the independent variable to K and writing

$$\varphi^2 = (m^2 + n^2)/m^2 A^2, \quad (41)$$

this equation becomes

$$\frac{\partial^2 \psi}{\partial K^2} - \frac{2}{K} \cdot \frac{\partial \psi}{\partial K} - \varphi^2 \psi = 0, \quad (42)$$

whose general solution is

$$\psi = C_1(\varphi K + 1)e^{-\varphi K} + C_2(\varphi K - 1)e^{\varphi K}, \quad (43)$$

C_1 and C_2 being arbitrary constants. By defining

$$\kappa = \sqrt{m^2 + n^2}, \quad (44)$$

and

$$\gamma = \kappa \cdot \frac{nB}{mA}, \quad (45)$$

the solution may be expressed in the form

$$\psi = C_1(\kappa x + \gamma + 1)e^{-(\kappa x + \gamma)} + C_2(\kappa x + \gamma - 1)e^{\kappa x + \gamma}. \quad (46)$$

Outside the current layer, where K takes the values defined by (21), ψ is governed by the equation

$$\frac{\partial^2 \psi}{\partial x^2} - \kappa^2 \psi = 0, \quad (47)$$

which has the solutions

$$\left. \begin{aligned} \psi &= C_3 e^{-\kappa x}, & x > a, \\ &= C_4 e^{\kappa x}, & x < -a, \end{aligned} \right\} \quad (48)$$

finite as $x \rightarrow \pm \infty$ respectively. C_3 and C_4 are arbitrary constants.

With the aid of (39) the corresponding solutions for h_x are found to be

$$\left. \begin{aligned} h_x &= \frac{i\kappa}{K}(\kappa x + \gamma) \{C_1 e^{-(\kappa x + \gamma)} - C_2 e^{\kappa x + \gamma}\}, & -a < x < a, \\ &= \frac{i\kappa}{K_+} C_3 e^{-\kappa x}, & x > a, \\ &= -\frac{i\kappa}{K_-} C_4 e^{\kappa x}, & x < -a, \end{aligned} \right\} \quad (49)$$

The requirement that h_x/K and ψ be continuous across the planes $x = \pm a$ imposes four relations on the constants C_1, \dots, C_4 , and it may be shown that the condition for the existence of non-trivial solutions is satisfied for all finite values of κa such that

$$e^{4\kappa a} = \frac{\{(\kappa a + \gamma + 1) - (\kappa a + \gamma)\alpha^2\}\{(\kappa a - \gamma + 1) + (\kappa a - \gamma)\beta^2\}}{\{(\kappa a + \gamma - 1) + (\kappa a + \gamma)\alpha^2\}\{(\kappa a - \gamma - 1) - (\kappa a - \gamma)\beta^2\}} \quad (50)$$

where

$$\left. \begin{aligned} \alpha &= \frac{mA_+ + nB_+}{mAa + nB}, \\ \beta &= \frac{mA_- + nB_-}{-mAa + nB} \end{aligned} \right\} \dots\dots\dots (51)$$

The condition (50) determines the stability of the current layer in the general case where the magnetic field in the fluid has the form specified by (19) and (21). The effect of the magnetic field on the stability enters into the condition (50) through the parameters α , β , and γ . In the following sections the implications of (50) are illustrated by reference to two special cases.

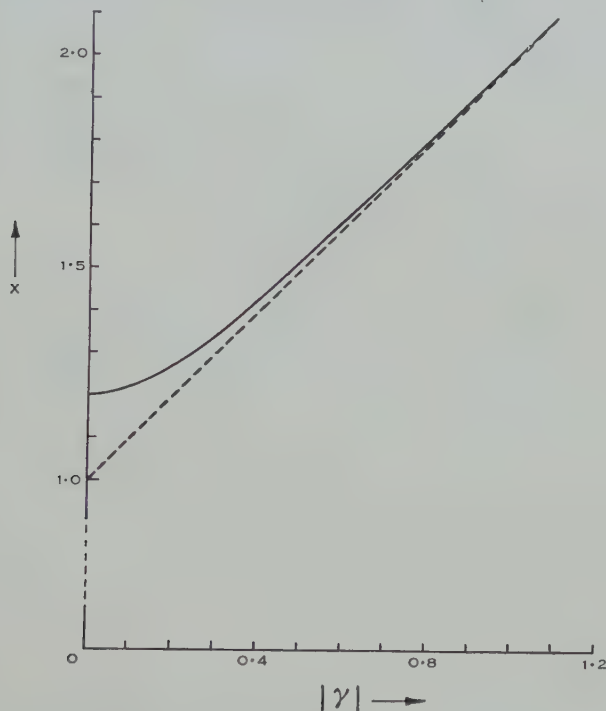


Fig. 2.—Variation of real positive root of equation (52) with $|\gamma|$.

V. ZERO MAGNETIC FIELD OUTSIDE THE CURRENT

If the magnetic field vanishes everywhere outside the region bounded by the planes $x = \pm a$, the parameters α and β in (50) are both zero and the stability condition assumes the limiting form

$$e^{4\gamma a} = \frac{(\gamma a + 1)^2 - \gamma^2}{(\gamma a - 1)^2 - \gamma^2} \dots\dots\dots (52)$$

It may be shown that this equation has just one positive real root $\gamma a = X$, say, corresponding to each value of $|\gamma|$. The variation of X with $|\gamma|$ is sketched in Figure 2. It is clear that X always exceeds $1 + |\gamma|$, but converges rapidly to this value with increasing $|\gamma|$.

To an excellent approximation, therefore, the solution of (52) may be written in the form

$$\kappa a = 1 + |\gamma|. \quad \dots\dots\dots (53)$$

Thus the relation (52) requires

$$a = \frac{1}{\sqrt{m^2 + n^2}} + \left| \frac{nB}{mA} \right| \quad \dots\dots\dots (54)$$

for instability. Since m and n may have any values between $-\infty$ and $+\infty$, there will clearly be a range of values of m and n for which any layer of finite thickness is unstable, and hence such a layer can have no permanent existence.



Fig. 3.—Variation of real positive root of equation (57) with ϵ .

VI. ZERO MAGNETIC FIELD IN DIRECTION OF CURRENT

In a recent paper Kruskal and Schwarzschild (1954) have examined the stability of a uniform current flowing through a region bounded by parallel planes, the magnetic field being taken to have no component in the direction of the current. On the assumption that the transverse field is greater outside the

region than within, these authors conclude that the system is neutral to perturbations in which the two defining surfaces are distorted into sinusoidal wave forms.

In this section the stability condition (50) is applied to the case where the transverse field outside the layer bears an arbitrary ratio to that inside. Suppose that the field inside is

$$H_y = \frac{\eta}{a} \cdot x, \quad \dots\dots\dots (55)$$

and the field outside

$$H_y = \pm \varepsilon \eta, \quad \dots\dots\dots (56)$$

where η is a constant and ε may take any positive value. In effect ε measures the ratio of the magnitude of the transverse field outside the region to that at the boundary of the current layer, and it is to be noted that the discussion is equally applicable when H_y has the same or opposite sign when $x > a$ and $x < -a$.

For this case the stability condition (50) assumes the form

$$e^{4\kappa a} = \frac{\{(1 + \varepsilon^2)\kappa a + 1\}\{(1 - \varepsilon^2)\kappa a + 1\}}{\{(1 + \varepsilon^2)\kappa a - 1\}\{(1 - \varepsilon^2)\kappa a - 1\}} \quad \dots\dots\dots (57)$$

The nature of the possible roots of equation (57) depends on the particular value given to the parameter ε . It may be shown that (57) has no real positive root for any $\varepsilon > 1$. A separate examination shows that for the value $\varepsilon = 1$ there is also no real positive root. Hence the current layer is stable for all $\varepsilon \geq 1$. This agrees with the result obtained by Kruskal and Schwarzschild.

However, when ε lies in the range $0 \leq \varepsilon < 1$, it may be shown that equation (57) has just one positive real root Y , say, for each value of ε . The variation of Y with ε is sketched in Figure 3. It is clear that Y always exceeds $1/(1 - \varepsilon^2)$, but converges to this value as $\varepsilon \rightarrow 1$. Hence to a good approximation, the solution of (57), may be written as

$$\kappa a = \frac{1}{1 - \varepsilon^2}, \quad \dots\dots\dots (58)$$

where $0 \leq \varepsilon < 1$. Equation (58) implies that, for any layer of finite thickness there are modes of disturbance for which the current distribution is unstable. Hence the state of the system passes from one of stability to one of instability when ε falls below unity.

VII. CONCLUSIONS

The general condition for the stability of a plane current layer has been derived by the method of normal modes and its implications have been discussed with reference to two special cases. In the first case the magnetic field has been taken to vanish everywhere outside the current layer, and it has been found that, for any layer of finite thickness, there are modes of disturbance for which the current distribution is unstable. The second case is an idealization of the well-known pinch effect, in which the current is self-constricted by a transverse magnetic field. It has been shown that the state of the system changes from one of stability to one of instability when the magnitude of the

field outside becomes less than that at the boundary of the current layer. This conclusion is in conformity with results obtained recently by Kruskal and Schwarzschild (1954).

The velocity of Alfvén waves along a magnetic field in a region bounded by parallel planes has been found to be reduced due to the inertia of the surrounding medium. This reduction is by a factor which approximates to $(1 + \mu_2/\mu_1)^{-\frac{1}{2}}$, where μ_1 is the mass density of the fluid inside the layer and μ_2 the density outside (cf. Fig. 1).

VIII. ACKNOWLEDGMENT

The author wishes to thank Dr. R. G. Giovanelli for his interest in this work.

IX. REFERENCES

- DUNGEY, J. W., and LOUGHHEAD, R. E. (1954).—*Aust. J. Phys.* **7**: 5.
 KRUSKAL, M., and SCHWARZCHILD, M. (1954).—*Proc. Roy. Soc. A* **223**: 348.

APPENDIX I

In Section II it was pointed out that the stability of the current layer does not depend on the compressibility or density of the conducting fluid. An equivalent result has been used in an earlier paper by Dr. J. W. Dungey and the author (1954), and it may be of some interest to give the proof in the case of the plane current layer.

For a compressible conducting fluid of variable density the equation describing small perturbations may be stated in the forms

$$\frac{\partial \mathbf{h}}{\partial t} = \text{curl} (\mathbf{v} \times \mathbf{H}), \quad \dots\dots\dots (59)$$

$$4\pi\mu\frac{\partial \mathbf{v}}{\partial t} = -\text{grad } \psi + (\mathbf{H} \cdot \text{grad})\mathbf{h} + (\mathbf{h} \cdot \text{grad})\mathbf{H}, \quad \dots\dots (60)$$

$$\frac{\partial \rho}{\partial t} = -\text{div} (\mu \mathbf{v}), \quad \dots\dots\dots (61)$$

where ρ is the variation in the mass density μ , which is now a function of x . Applying the method of normal modes, solutions are sought in which the variables \mathbf{h} , \mathbf{v} , ψ , and ρ are taken proportional to $\exp i(\omega t + my + nz)$. The conditions for instability are then obtained by putting $\omega = 0$; but in this case equations (59) and (61) require $\mathbf{v} = 0$, and so equation (60) must be supplemented by the divergence relation $\text{div } \mathbf{h} = 0$ to obtain a determinate solution.

Hence the equations determining the onset of instability in a compressible fluid of variable density are identical with those obtained previously for the case of an incompressible fluid of uniform density. The conditions for instability are thus unaltered when compressibility and non-uniform density are taken into account.

CORRECTING FOR RUNNING MEANS BY SUCCESSIVE SUBSTITUTIONS

By R. N. BRACEWELL*

[*Manuscript received April 22, 1955*]

Summary

The paper discusses a practical procedure which compensates for the effect of taking running means, and a numerical example is worked out. The procedure simply requires the evaluation of further running means and should prove readily applicable in many cases where the need arises. In this procedure, which has already gained importance in more general form in other fields (as the method of successive substitutions), a criterion of convergence has been given; and it is a major aim of the paper to illustrate, using running means as an instance, that the utility of the method of successive substitutions is wider than is indicated by this criterion. The mathematical theory shows that the proposed procedure leads to a divergent result in the case of running means; nevertheless the asymptotic nature of the divergence allows results of practical value. An illuminating view of the phenomenon is given from the standpoint of Fourier analysis, which reveals a counterplay of simultaneous deterioration and improvement occurring in different spectral regions.

I. INTRODUCTION

It often happens, when a quantity $f(x)$ has to be observed at $x=x_1$, that the quantity actually measured is a weighted integral of $f(x)$ over an interval surrounding x_1 . This interval can sometimes be made so small that further decrease makes no difference, but often this cannot be done, and the problem then arises of allowing for the non-zero width of the interval.

The nature of the weighting function varies with the circumstances. In spectroscopy it is the "apparatus profile", a more or less smooth humped curve; in radio astronomy it is the "aerial directional diagram", a central hump with "side lobes"; and there are numberless other cases. In many circumstances the weighting function is rectangular—this is the important case where the observation is a simple running mean of the desired quantity. An example of this occurs when a photographic density distribution $f(x)$ is scanned by a slit.

This paper deals with a method of correcting for running means. The method, which in its general form has been presented at length in connexion with the radio-astronomical problem by Bracewell and Roberts (1954), will be referred to as the method of successive substitutions as in the theory of integral equations (e.g. Lovitt 1950). The method seems to have been introduced into

* Division of Radiophysics, C.S.I.R.O., University Grounds, Sydney, and Berkeley Astronomical Department, University of California, Berkeley.

astronomy by van Cittert (1931), but is not well known in some branches of physics where it has immediate application (e.g. Bracewell 1955). It yields a sequence of distributions, each derived from its predecessor by a process involving a further application of the very effect which it is desired to eliminate. In the case of running means the procedure, briefly stated, would be as follows. Take running means of the observations, note the small change produced, and apply this change as a negative correction to the observations. This gives the second member of a sequence of distributions of which the observed distribution itself is the first member. The next distribution is obtained by taking running means of the second, comparing the result so obtained with the observations, and applying the discrepancy as a correction to the second distribution. If the discrepancy tends, in successive stages, towards zero, then the distributions are tending to a form which, when smoothed by running means, agrees more and more closely with the original observations. In other words, the sequence of distributions is converging to a desired solution.

This procedure involves nothing more recondite than repeated running means of the observed data, and, as running means are often easy to get, such a procedure would be very attractive in many applications.

A note of caution now needs to be sounded, for only in favourable cases does the sequence converge. For evidence of this, one need only ponder the case $f(x) = \delta(x)$ and weighting function $\frac{1}{2}\delta(x+1) + \frac{1}{2}\delta(x-1)$, where $\delta(x)$ is the unit impulse function.

Now the criterion of convergence is known and, when applied to running means, reveals that the sequence converges only for initial distributions with a certain peculiar type of spectrum. An understanding of the convergence question is therefore all important in this method of correcting for running means. On the other hand, in radio astronomy, convergence can normally be taken for granted.

The method of successive substitutions is, however, likely to prove of far more general use than its failure to satisfy convergence criteria might lead one to suppose. As explained later in connexion with Figure 1, it may happen that the sequence of distributions, whilst ultimately divergent, nevertheless for a few stages approaches the desired solution asymptotically. As it is probable that, in the bulk of applications, a procedure would not be considered practicable which did not give adequate correction in the first one or two stages, it should be immaterial in practice whether the sequence is, in the end, convergent or divergent. This paper is mainly concerned with showing that excellent results may often be had in spite of divergence, using the case of running means for the purpose.

II. DEFINITION OF THE PROBLEM

Let $g(x)$ be derived from a function $f(x)$ by averaging over an interval ξ . Then

$$g(x) = \xi^{-1} \int_{x-\frac{1}{2}\xi}^{x+\frac{1}{2}\xi} f(\tau) d\tau. \quad \dots\dots\dots (1)$$

The problem is, given $g(x)$, to find $f(x)$.

Rewrite equation (1) in the form

$$g(x) = \xi^{-1} \int_{-\infty}^{\infty} \Pi\left(\frac{x-\tau}{\xi}\right) f(\tau) d\tau, \quad \dots\dots\dots (2)$$

where $\Pi(\tau)$ is the rectangle function of unit height and breadth :

$$\begin{aligned} \Pi(\tau) &= 1, & -\frac{1}{2} < \tau < \frac{1}{2}, \\ &= 0, & |\tau| > \frac{1}{2}. \end{aligned}$$

The right-hand side of equation (2) may be recognized as a convolution integral, which, for clarity, may be written with the asterisk notation† as

$$g(x) = \Pi_{\xi} * f, \quad \dots\dots\dots (3)$$

where $\Pi_{\xi}(x) = \xi^{-1} \Pi(\xi^{-1}x)$.

Equation (3) is strictly analogous to the aerial smoothing equation discussed by Bracewell and Roberts, and it will be seen that Π_{ξ} is already correctly normalized in the sense that

$$\int_{-\infty}^{\infty} \Pi_{\xi}(x) dx = 1.$$

The sequence of distributions yielded by the method of successive substitutions is therefore as follows :

$$\left. \begin{aligned} g, \\ f_1 &= 2g - \Pi_{\xi} * g, \\ f_2 &= 3g - 3\Pi_{\xi} * g + \Pi_{\xi} * \Pi_{\xi} * g, \\ &\dots\dots\dots \\ &\dots\dots\dots \end{aligned} \right\} \dots\dots\dots (3a)$$

It has now to be considered whether the members of this sequence approach the desired distribution f , or, if not, whether any of them is an improvement on g .

III. CONDITION FOR CONVERGENCE

According to Bracewell and Roberts, the sequence converges if, and only if,

$$|1 - \overline{\Pi}_{\xi}(s)| < 1$$

for all s for which $\bar{g}(s) \neq 0$, where $\overline{\Pi}_{\xi}(s)$ and $\bar{g}(s)$ are respectively the Fourier transforms of Π_{ξ} and g . In general the sequence converges not to f , but to the function whose transform is

$$\frac{\bar{f}(s)}{1 + \delta(\overline{\Pi}_{\xi})}.$$

Now

$$\overline{\Pi}_{\xi}(s) = \frac{\sin \pi \xi s}{\pi \xi s}. \quad \dots\dots\dots (4)$$

† In this notation $f * g \equiv \int_{-\infty}^{\infty} f(x-u)g(u)du$.

Hence $|1 - \bar{\Pi}_\xi(s)| \geq 1$ for all s such that

$$n\xi^{-1} \leq s \leq (n+1)\xi^{-1},$$

where n is any odd integer.

Consequently the sequence converges only in the cases where $g(x)$ has no spectral components in the shaded zones of Figure 1.

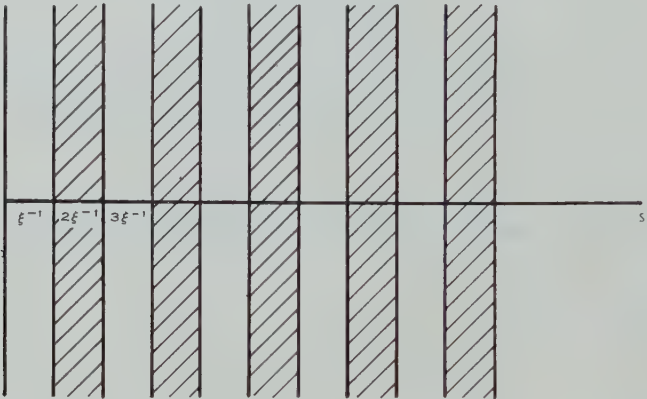


Fig. 1

Few distribution functions in practice could be expected to comply with such a severe condition. Therefore, in general, the sequence could not be relied on to converge. However, there remains the possibility of asymptotic representation and this may be illustrated with a numerical example.

We adopt the symmetrical distribution $g(x)$, as given in Table 1 and graphed in Figure 2, constructed from the assumed distribution $f(x)$ by taking running

TABLE 1
AN EXAMPLE OF ONE STAGE OF CORRECTION

x	$f(x)$	$g(x)$	$\Pi\xi * g$	f_1	Discrepancy
0	8800	8400	8032	8768	-32 (½%)
	8600	8200	7880	8520	-80 (1%)
2	8000	7680	7440	7920	-80 (1%)
	7000	6920	6760	7080	80 (1%)
4	6000	6000	5920	6080	80 (1%)
	5000	5000	4992	5008	8 (2%)
6	4000	4000	4040	3960	-40 (1%)
	3000	3040	3120	2960	-40 (1%)
8	2000	2160	2280	2040	40 (2%)
	1200	1400	1560	1240	40 (3%)
10	600	800	984	616	16 (3%)
	200	400	560	240	40 (20%)
12	0	160	280	40	40
	0	40	120	-40	-40
14	0	0	40	-40	-40
	0	0	8	-8	-8

means over five values. The column $\Pi_{\xi} * g$ is derived from $g(x)$ in the same way that $g(x)$ is derived from $f(x)$. The column $f_1 = 2g - \Pi_{\xi} * g$ gives the result of one stage of correction, and the last column gives the discrepancy between the corrected distribution $f_1(x)$ and the known original distribution $f(x)$.

Figure 2 shows the agreement between the corrected distribution (-----) and the known original function $f(x)$, an agreement which for many practical purposes would clearly be very acceptable.

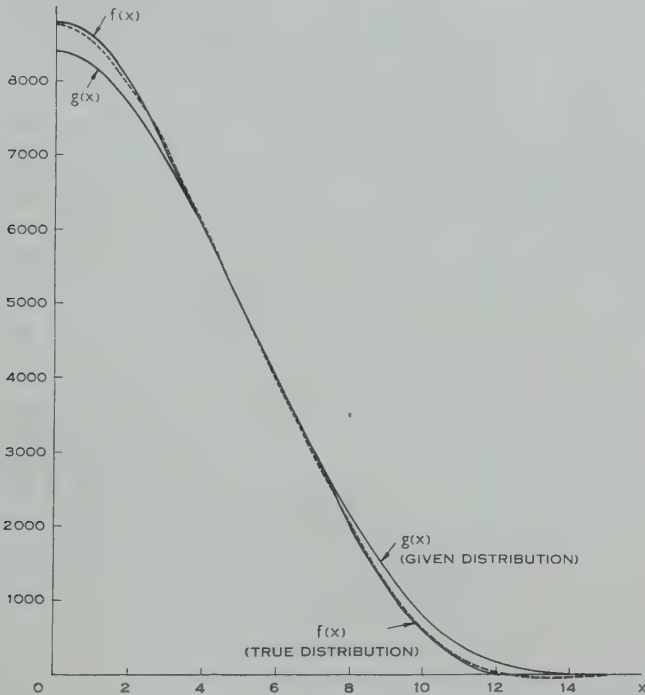


Fig. 2

IV. SPECTRAL ASPECT OF ASYMPTOTIC PHENOMENA

Applying the convolution theorem to (3) we have

$$\bar{g} = \bar{\Pi}_{\xi} \bar{f}.$$

Hence, from (4),

$$\bar{g} = \frac{\sin \pi \xi s}{\pi \xi s} \bar{f}.$$

This relationship between the spectra \bar{f} and \bar{g} is illustrated qualitatively in Figure 3. Now consideration of equation (3a) will show that the transforms of the successive distributions fall as in the broken line of Figure 3. In the non-shaded strips, the broken line is closer to \bar{f} than \bar{g} is, but in the shaded strips it is not, and tends to depart further with succeeding stages.

If then a stage of restoration is to improve the agreement with f , the improvement occurring in the unshaded strips must outweigh the deterioration setting

in in the shaded strips. This will occur when the spectrum of f is weak in the shaded strips, as may be shown to be the case in the example of Figure 2.

V. CRITERIA FOR PRACTICAL USE

We have now seen that satisfactory results can follow from asymptotic sequences, and an illuminating explanation of what happens has been possible. In practice, however, the results cannot be verified in the way used here, nor can Fourier transforms be considered in detail. A criterion is therefore required by which it can be decided whether the corrected distribution $f_1(x)$ represents an improvement over $g(x)$ as an approximation to the solution $f(x)$.

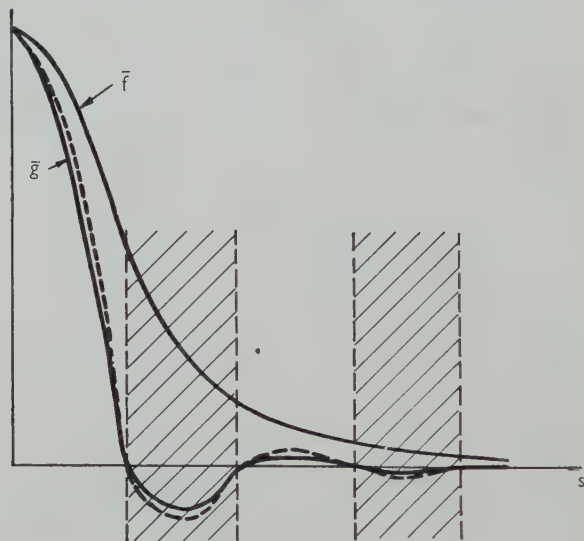


Fig. 3

Fortunately the method itself gives an indication, for, in cases of convergence, successive corrections become smaller and smaller. Now it is not reasonable to attempt to assess the quality of a distribution numerically. For example, a corrected distribution might show improvement in some respects (say general shape) and deterioration in others (presence of ripple); and in some circumstances this might be on the whole "better" and in other circumstances "worse". The required criterion must therefore be a qualitative one involving the user's needs; thus as a practical test, it would seem to be sufficient that $\Pi\xi*f_1$ should agree with g much "better" (for the existing purpose) than g does with f_1 .

VI. REFERENCES

- BRACEWELL, R. N. (1955).—A method of correcting the broadening of X-ray line profiles. *Aust. J. Phys.* **8**: 61.
 BRACEWELL, R. N., and ROBERTS, J. A. (1954).—Aerial smoothing in radio astronomy. *Aust. J. Phys.* **7**: 615.
 VAN CITTERT, P. H. (1931).—Zum Einfluss der Spaltbreite auf die Intensitätsverteilung in Spektrallinien. *Z. Phys.* **69**: 298.
 LOVITT, W. V. (1950).—"Linear Integral Equations." (Dover: New York.)

THE HIGH TEMPERATURE EXCITATION OF HELIUM

By J. T. JEFFERIES*

[*Manuscript received April 4, 1955*]

Summary

The relative populations of the helium ground state, the singlet and triplet states of principal quantum number 2 and 3, and the ionized state are obtained for an atmosphere isothermal at one of a number of kinetic temperatures in the range 10^4 – 10^5 °K.

It is shown that the non-appearance of the D_3 line of helium in the solar Fraunhofer spectrum may be reconciled with the simultaneous appearance of the infra-red $\lambda 10830$ line, provided the mean kinetic temperature of the chromosphere is not too great. From an analysis of observations of the D_3 emission of the solar chromosphere, it is concluded that the mean kinetic temperature of the lower chromosphere is most probably in the range 10^4 – 1.25×10^4 °K.

I. INTRODUCTION

The suggestion by Redman (1942) that the solar chromosphere has a high kinetic temperature ($\sim 30,000$ °K) has in recent years stimulated studies of the excitation of and emission of radiation from high temperature atmospheres, with particular reference to the solar chromosphere and various related features such as prominences and flares. Attention so far has been confined largely to the excitation of hydrogen since it is the most abundant element in stellar systems and the simplest to consider theoretically and because most of the relevant atomic excitation processes have been studied. The data so obtained have been applied by the author (1955) to determine the physical conditions in prominences, and it is clearly desirable that such computations should be extended to other elements so as to check results found for hydrogen and to allow interpretation of observations where the results for hydrogen are either lacking or equivocal.

In this paper the excitation of helium is considered for a number of kinetic temperatures in the range 10^4 – 10^5 °K, this range being of considerable interest in astrophysics. The choice of helium as the next element for consideration after hydrogen lies in its importance in the solar flash spectrum and its high excitation potential which should make the emission quite sensitive to small changes in kinetic temperature. Further, the interpretation of helium spectra from the Sun is likely to be more straightforward than for hydrogen, as the effects of self-absorption should be far smaller. The rather curious and variable behaviour of helium eclipse spectra also makes its study of interest, as it is felt that an explanation of such features could assist considerably in solving problems of chromospheric structure.

* Division of Physics, C.S.I.R.O., University Grounds, Sydney.

In order to test Redman's (1942) suggestion of a high chromospheric temperature, Miyamoto (1951) considered the excitation of helium in a model solar chromosphere at 35,000 °K and showed that, at such a high temperature, one would expect strong He I and He II absorption lines in the solar disk spectrum, contrary to observation.

We shall consider here the excitation of helium in a high temperature, horizontally stratified, isothermal atmosphere of which helium is a constituent. The presence of other elements in the atmosphere is taken to influence the helium excitation only by increasing the electron concentration and so increasing the rate of collisional excitation; heavy particle collisions are assumed of negligible importance. In general the most important element present, at least in astrophysical situations, is hydrogen, and the numerical illustrations given here are limited to cases where the atmosphere is composed of hydrogen and helium, generally with an abundance ratio 5:1. In these cases, the electrons result mainly from ionization of hydrogen. The theory, however, is developed using the electron concentration N_e as a parameter and so may be applied to any atmosphere.

Excitations and de-excitations of the atoms are also caused by radiative transitions involving absorption or emission of radiation. Since applications will be mainly to solar phenomena, numerical results are confined to the case of an atmosphere irradiated by hemispherical black-body radiation at 6000 °K, corresponding to solar photospheric radiation. This irradiation will only be significant in those parts of the spectrum where the optical depth of the atmosphere is small. Elsewhere radiative excitation is due to the radiation field built up by the atmosphere itself.

As for hydrogen, which was treated by Jefferies and Giovanelli (1954),* the problem for helium may be formulated in terms of a set of simultaneous equilibrium equations which state that the rate of entry into an atomic level equals the rate of leaving, together with a set of second order differential equations whose solutions give the intensities in each spectral line and continuum. As the solution of such a set is formidable, recourse is made to an approximate treatment using a restricted number of energy levels, namely, the ground and ionized states together with the singlet and triplet states of principal quantum numbers 2 and 3. The fine structure splitting will be ignored; where necessary we shall suppose these states to be populated in the ratios of their statistical weights.

Intercombination transitions between an excited singlet and a triplet level will also be ignored. Optically forbidden radiative transitions of this type are known to be weak in comparison with the corresponding optically allowed transitions. Collisional transitions will be neglected for simplicity, but this neglect will not cause much error as the singlet and triplet levels are quite strongly linked through the ionized state. Further these intercombination collision transitions are unlikely to be more frequent than the transitions between states of equal multiplicity, and, since we can hardly do better than guess these latter rates, the order of accuracy of the results will not be greatly affected by the

* This paper will be referred to throughout as paper A.

neglect of intercombination transitions. Transitions between the ground state and the triplet 2 and 3 states, whose cross sections are quite well known, are, however, taken into account.

II. THE EXCITATION RATES

The total rate of excitation per unit volume from state j to state l is denoted $P_{jl}N_j$ where N_j is the population of the j state. P_{jl} may be dissected into a radiative component A_{jl} and an electron collisional component R_{jl} which contains the electron concentration N_e as a factor. The rates of collision excitation

TABLE 1
HELIUM COLLISION CROSS SECTIONS (πa_0^2)

Transition	Cross Section
1^1S-2^1S	$3 \times 10^{-2} (E_0/E)$
1^1S-2^1P	$1.5 \times 10^{-1} (E-E_0)/(E_m-E_0); E_m=50 \text{ eV}$
1^1S-3^1S	$3 \times 10^{-3} (E_0/E)$
1^1S-3^1P	$3 \times 10^{-2} (E-E_0)/(E_m-E_0); E_m=100 \text{ eV}$
1^1S-3^1D	$8 \times 10^{-4} E_0/E$
1^1S-2^3S	$4 \times 10^{-2} (E_0/E)^2$
1^1S-2^3P	$3 \times 10^{-2} (E_0/E)^3$
1^1S-3^3S	$2 \times 10^{-3} (E_0/E)^2$
1^1S-3^3P	$1 \times 10^{-2} (E_0/E)^3$
1^1S-3^3D	$8 \times 10^{-4} (E_0/E)^4$

follow, assuming a Boltzmann distribution of energy among the electrons, from the absolute values and energy dependence of the collision cross sections. For collisions from the ground state this information is available for most transitions—Bates *et al.* (1950)—and the values adopted here are shown in Table 1 together with our idealized energy dependence. For the optically allowed transitions 1^1S-2^1P , 3^1P , a linear energy dependence is assumed here up to a maximum

TABLE 2
ADOPTED VALUES OF q

Transition	$2S-3S$	$2S-3P$	$2S-3D$	$2P-3S$	$2P-3P$	$2P-3D$
q (eV)	10	100	5	10	10	150

cross section at E_m . Owing to the high values of E_m compared with kT , the value of the excitation rate is effectively independent of the cross section beyond E_m . For transitions between excited states some reasonable guess must be made as no data are available. It has been assumed that, for corresponding singlet-singlet and triplet-triplet transitions, the cross sections are equal and of the form $q\pi a_0^2/E$; the adopted values of q obtained from comparison with corresponding data for hydrogen are shown in Table 2.

The resultant collisional excitation rates between states of given principal quantum numbers are shown in Table 3 for a number of kinetic temperatures.

The rates of super-elastic collision and of collision ionization are found in a straightforward manner by the method used by Giovanelli (1948). The rates of photoelectric ionization and of spontaneous recombination are also computed

TABLE 3
RATES OF COLLISION EXCITATION ($N_e N_n$ CM⁻³ SEC⁻¹)

T (10 ⁴ °K)	1-2 ¹	1-3 ¹	2-3	1-2 ³	1-3 ³
1.00	3.5×10^{-19}	2.2×10^{-21}	1.0×10^{-7}	7.5×10^{-19}	5.3×10^{-21}
1.25	4.2×10^{-17}	3.4×10^{-19}	1.4×10^{-7}	6.5×10^{-17}	9.0×10^{-19}
1.50	1.0×10^{-15}	1.0×10^{-17}	1.8×10^{-7}	1.3×10^{-15}	2.6×10^{-17}
2.00	6.3×10^{-14}	8.4×10^{-16}	2.2×10^{-7}	5.6×10^{-14}	1.8×10^{-15}
2.50	7.2×10^{-13}	1.0×10^{-14}	2.5×10^{-7}	4.9×10^{-13}	2.2×10^{-14}
5.00	1.2×10^{-10}	1.8×10^{-12}	2.8×10^{-7}	3.7×10^{-11}	3.3×10^{-12}
10.00	1.8×10^{-9}	2.4×10^{-11}	2.4×10^{-7}	2.7×10^{-10}	3.2×10^{-11}

by a similar method using the Gaunt formula for the continuous absorption coefficient, as modified by Menzel (1931).

$$\alpha(n, \nu) = 2^6 \pi^4 m Z^4 \epsilon^{10} / 3 \sqrt{3} c h^6 \nu^3 n_{\text{eff}}^5,$$

where n_{eff} is the effective principal quantum number for the level.

Rates of spontaneous transitions between bound states may be computed from tables given by Bates and Damgaard (1949) and are shown in Table 4.

TABLE 4
SPONTANEOUS TRANSITION RATES (N_j CM⁻³ SEC⁻¹)

Transition $j-l$..	2 ¹ -1	3 ¹ -1	3 ¹ -2 ¹	3 ³ -2 ³
Rate	1.62×10^9	2.15×10^8	4.27×10^7	4.60×10^7

Of the intercombination lines, one would expect the most important to be the extreme ultraviolet line resulting from the transition 2³P-1¹S. Its transition probability, however, seems to be very small, as can be inferred from photographs given by Hopfield (1930) and we shall ignore it here in comparison with other processes removing atoms from the triplet 2 state. It should, however, be pointed out that this intercombination line is likely to be important in very diffuse atmospheres only faintly illuminated from outside, since in such cases it is likely to be the predominant transition from the triplet 2 state.

III. THE POPULATIONS OF THE GROUND AND IONIZED STATES

Using excitation rates found as in Section II, we may formulate six equilibrium equations for the populations of the 1, 2¹, 3¹, 2³, 3³, and 4 states, the last representing the ionized state, He II. This set may be solved for the ratios of populations of any two states; for absolute values we need, of course, one

further piece of information which will usually be the total concentration of helium atoms and ions.

The Lyman type radiation arising from the transitions 2^1-1 , 3^1-1 , and $4-1$ will be denoted here by $L\alpha$, $L\beta$, and $L\gamma$. Since we shall be dealing solely with helium radiation, no confusion with the Lyman radiation of hydrogen should result.

Arrivals in the ionized, 4, state are due to :

- (i) ionizations from 1,
- (ii) ionizations from 2^1 and 3^1 ,
- (iii) ionizations from 2^3 and 3^3 .

If the atmosphere is optically thin to $L\alpha$ and $L\beta$, the contributions from (ii) are negligible compared with (i). Ionizations from the triplet states, however, are not negligible since, following the collision excitations $1-2^3$, $1-3^3$ the most probable result is an ionization, since the rates of the super-elastic collision 3^3-1 , 2^3-1 are much less than the corresponding rates of ionization. The ionizations (i) and (iii) are balanced by recombinations to the ground and excited singlet levels ; any recombination to the triplet levels resulting in a subsequent ionization. It follows then that

$$(R_{14} + R_{12^3} + R_{13^3})N_1 = (P_{41} + P_{42^1} + P_{43^1})N_4,$$

or

$$\frac{N_1}{N_4} = \frac{P_{41} + P_{42^1} + P_{43^1}}{R_{14} + R_{12^3} + R_{13^3}} \dots \dots \dots (3.1)$$

Excited singlet levels have a high probability of returning direct to the ground state emitting a quantum of radiation. If, however, the atmosphere is optically thick to $L\alpha$ and $L\beta$, this quantum is immediately reabsorbed in the atmosphere and the process is repeated until an ionization occurs. Thus nearly all collisions from the ground state result eventually in ionizations, which together with the ionizations from the ground state are balanced by recombinations to the ground state only, since a recombination to any excited level will result in a subsequent ionization in the great majority of cases. For high opacity in $L\alpha$ and $L\beta$ therefore

$$\frac{N_1}{N_4} = \frac{P_{41}}{P_{14} + R_{12^1} + R_{13^1} + R_{12^3} + R_{13^3}} \dots \dots \dots (3.2)$$

If the atmosphere is optically thin in $L\gamma$, P_{14} may be replaced by R_{14} and hence

$$\frac{N_1}{N_4} = \frac{P_{41}}{\sum_l R_{1l}} \dots \dots \dots (3.3)$$

The values of N_1/N_4 shown in Table 5 are obtained on inserting numerical values into (3.1) and (3.3).

If the optical thickness in $L\gamma$ is not small, P_{14} in (3.2) cannot be replaced by R_{14} , i.e. account must be taken of the effect of $L\gamma$ absorption on the populations N_1 and N_4 . Applying the method used in paper A for the corresponding problem for hydrogen it is found that, for high opacity in $L\alpha$, $L\beta$, and $L\gamma$,

$$\frac{N_1}{N_4} = \lambda_{14} \frac{P_{41}}{\sum_l R_{1l}}, \dots \dots \dots (3.4)$$

TABLE 5
THE RATIO N_1/N_4

T (10^4 °K)	Atmospheres Thin in $L\alpha$, $L\beta$, $L\epsilon$ Eqn. (3.1)	Atmospheres Thick in $L\alpha$, $L\beta$ and Thin in $L\epsilon$ Eqn. (3.3)
1.00	4.3×10^5	2.4×10^5
1.25	4.4×10^3	2.2×10^3
1.50	1.9×10^2	8.8×10^1
2.00	3.1	1.4
2.50	2.6×10^{-1}	1.1×10^{-1}
5.00	9.3×10^{-4}	4.5×10^{-4}
10.00	5.0×10^{-5}	2.2×10^{-5}

where λ_{14} is the scattering parameter for $L\epsilon$ radiation, whose value is found, using the general method given by Giovanelli and Jefferies (1954), to be given approximately by the relation

$$\lambda_{14} = \frac{\alpha'(P_{42^3} + P_{43^3}) + \alpha(P_{42^1} + P_{43^1})}{P_{41}}, \quad \dots\dots\dots (3.5)$$

where $\alpha' = R_{2^1 1}/P_{2^3 4}$ and $\alpha = R_{2^1 1}/P_{2^1 4}$.

Expressions for λ_{14} in terms of N_e are shown in Table 6.

TABLE 6
THE SCATTERING PARAMETER FOR $L\epsilon$ RADIATION

T (10^4 °K)	λ_{14}
1.00	$7.6 \times 10^{-14} N_e / (1 + 3.7 \times 10^{-13} N_e)$
1.25	$6.4 \times 10^{-14} N_e / (1 + 1.0 \times 10^{-12} N_e)$
1.50	$6.3 \times 10^{-14} N_e / (1 + 1.6 \times 10^{-12} N_e)$
2.00	$5.2 \times 10^{-14} N_e / (1 + 3.3 \times 10^{-12} N_e)$
2.50	$4.8 \times 10^{-14} N_e / (1 + 5.0 \times 10^{-12} N_e)$
5.00	$3.8 \times 10^{-14} N_e / (1 + 1.1 \times 10^{-11} N_e)$
10.00	$3.4 \times 10^{-14} N_e / (1 + 1.5 \times 10^{-11} N_e)$

Optical depths for an isothermal atmosphere of temperature T are given below. The values refer to the line centres for $L\alpha$ and $L\beta$ and to the head of the continuum for $L\epsilon$.

$$\tau_{L\alpha} = \frac{4.5 \times 10^{-12}}{\sqrt{T}} \int_0^\infty N_1 dz,$$

$$\tau_{L\beta} = \frac{4.7 \times 10^{-13}}{\sqrt{T}} \int_0^\infty N_1 dz,$$

$$\tau_{L\epsilon} = 6.2 \times 10^{-18} \int_0^\infty N_1 dz.$$

Thus for a model atmosphere of thickness 2×10^8 cm, electron concentration $5 \times 10^{11} \text{ cm}^{-3}$, and kinetic temperature $T = 10^4$ °K in which the H : He abundance ratio is 5 we find, since the hydrogen is nearly all ionized—paper A—while the helium atoms are nearly all in the ground state, that $N_1 \sim 10^{11} \text{ cm}^{-3}$, and

$$\tau_{L\alpha} \sim 9 \times 10^5,$$

$$\tau_{L\beta} \sim 9 \times 10^4,$$

$$\tau_{Lc} \sim 10^2,$$

while for the same atmosphere at 5×10^4 °K, $N_1 \sim 5 \times 10^7$ and

$$\tau_{L\alpha} \sim 2 \times 10^2,$$

$$\tau_{L\beta} \sim 2 \times 10^1,$$

$$\tau_{Lc} \sim 6 \times 10^{-2}.$$

IV. THE POPULATIONS OF THE EXCITED STATES

The equilibrium equations for the excited singlet or triplet states may be written

$$\left. \begin{aligned} P_{42}N_4 + P_{32}N_3 + P_{12}N_1 &= T_2N_2, \\ P_{43}N_4 + P_{23}N_2 + P_{13}N_1 &= T_3N_3, \end{aligned} \right\} \dots\dots\dots (4.1)$$

where T_2 and T_3 represent the total rates per atom for departure from the 2 and 3 states and, since the equations apply to either singlet or triplet levels, distinguishing symbols have been omitted.

The solutions of (4.1) are

$$\left. \begin{aligned} N_2 &= \frac{T_3K_2 + P_{32}K_3}{T_2T_3 - P_{23}P_{32}}, \\ N_3 &= \frac{T_2K_3 + P_{23}K_2}{T_2T_3 - P_{23}P_{32}}, \end{aligned} \right\} \dots\dots\dots (4.2)$$

where

$$K_2 = P_{42}N_4 + P_{12}N_1,$$

$$K_3 = P_{43}N_4 + P_{13}N_1.$$

If the atmosphere is optically thin in the 2-3 radiation, the radiative component A_{23} of the transition rate P_{23} is found from the intensity of the radiation incident on the atmosphere. If it is not optically thin, A_{23} is found from the intensity of the radiation inside the atmosphere, which varies from depth to depth and may be computed in the manner indicated in paper A. The following remarks apply to an atmosphere optically thin in 2-3 radiation, although the modifications necessary for a thick atmosphere are straightforward.

For high opacity in the Lyman type lines, we replace the transition rates P_{12^1} , $P_{2^1 1}$, P_{13^1} , $P_{3^1 1}$ by their corresponding collisional transition rates and obtain, neglecting some small terms, and assuming $P_{23} > P_{24}$, which will be so for cases of interest here,

$$\left. \begin{aligned} N_{2^1} &\simeq \frac{K_{2^1} + K_{3^1}}{P_{2^1 4}}, \\ N_{3^1} &\simeq \frac{P_{2^1 3^1}}{P_{3^1 2^1}} N_{2^1}. \end{aligned} \right\} \dots\dots\dots (4.3)$$

The triplet populations N_{2^3} and N_{3^3} are also given by equations of the form (4.3) owing to the assumption of no radiative intercombination between the ground and triplet levels.

If the atmosphere is optically very thin in $L\alpha$ and $L\beta$ solutions (4.2) become, for the singlet states,

$$\left. \begin{aligned} N_{2^1} &\simeq \frac{R_{12^1}N_1 + P_{42^1}N_4}{P_{2^1}}, \\ N_{3^1} &\simeq \frac{R_{13^1}N_1 + P_{43^1}N_4}{P_{3^1}}, \end{aligned} \right\} \dots\dots\dots (4.4)$$

TABLE 7
POPULATIONS OF THE SINGLET AND TRIPLET 2 STATES

T (10^4 °K)	N_{2^1}/N_1N_e (cm^3)	N_{2^3}/N_1N_e (cm^3)
1.00	$\frac{3.5 \times 10^{-19} + 7.2 \times 10^{-14}N_4/N_1}{2.5 \times 10^4 + 8.8 \times 10^{-9}N_e}$	$\frac{7.5 \times 10^{-19} + 2.4 \times 10^{-13}N_4/N_1}{1.4 \times 10^4 + 5.5 \times 10^{-9}N_e}$
1.25	$\frac{4.2 \times 10^{-17} + 6.1 \times 10^{-14}N_4/N_1}{2.5 \times 10^4 + 2.0 \times 10^{-8}N_e}$	$\frac{6.6 \times 10^{-17} + 2.0 \times 10^{-13}N_4/N_1}{1.4 \times 10^4 + 1.4 \times 10^{-8}N_e}$
1.50	$\frac{1.0 \times 10^{-15} + 5.3 \times 10^{-14}N_4/N_1}{2.5 \times 10^4 + 3.5 \times 10^{-8}N_e}$	$\frac{1.3 \times 10^{-15} + 1.8 \times 10^{-13}N_4/N_1}{1.4 \times 10^4 + 2.5 \times 10^{-8}N_e}$
2.00	$\frac{6.4 \times 10^{-14} + 4.3 \times 10^{-14}N_4/N_1}{2.5 \times 10^4 + 7.3 \times 10^{-8}N_e}$	$\frac{5.8 \times 10^{-14} + 1.4 \times 10^{-13}N_4/N_1}{1.4 \times 10^4 + 5.3 \times 10^{-8}N_e}$
2.50	$\frac{7.3 \times 10^{-13} + 3.4 \times 10^{-14}N_4/N_1}{2.5 \times 10^4 + 1.1 \times 10^{-7}N_e}$	$\frac{5.1 \times 10^{-13} + 1.2 \times 10^{-13}N_4/N_1}{1.4 \times 10^4 + 8.2 \times 10^{-8}N_e}$
5.00	$\frac{1.2 \times 10^{-10} + 2.0 \times 10^{-14}N_4/N_1}{2.5 \times 10^4 + 2.4 \times 10^{-7}N_e}$	$\frac{4.0 \times 10^{-11} + 6.9 \times 10^{-14}N_4/N_1}{1.4 \times 10^4 + 2.0 \times 10^{-7}N_e}$
10.00	$\frac{1.8 \times 10^{-9} + 1.1 \times 10^{-14}N_4/N_1}{2.5 \times 10^4 + 3.4 \times 10^{-7}N_e}$	$\frac{3.0 \times 10^{-10} + 3.7 \times 10^{-14}N_4/N_1}{1.4 \times 10^4 + 2.8 \times 10^{-7}N_e}$

which are much smaller than the values given by (4.3). For intermediate optical depths in $L\alpha$ and $L\beta$ approximations to the ratios N_{2^1}/N_1 and N_{3^1}/N_1 may be found by the method indicated in paper A. We outline the method for N_{2^1}/N_1 ; the procedure for N_{3^1}/N_1 is analogous.

Provided the optical depth in $L\alpha$ is not too low,

$$\frac{N_{2^1}}{N_1} \simeq \frac{A_{12^1}}{A_{21^1}} \simeq \frac{J_\alpha}{\rho_\alpha}, \dots\dots\dots (4.5)$$

where $\rho_\alpha = \frac{8\pi h\nu_\alpha^3}{c^2} \frac{\omega_1}{\omega_2}$, ω_1 and ω_2 are the statistical weights, and J_α is the total intensity at the level in the atmosphere where the ratio N_2/N_1 is required. The value of J_α is found as in paper A; for a homogeneous atmosphere with no incident radiation at the top and total reflection at the base,

$$J = \frac{4\pi b}{\lambda} \left[1 - \frac{\exp(-\sqrt{3\lambda}\tau) + \exp[\sqrt{3\lambda}(\tau - 2\tau_1)]}{1 + 2\sqrt{\lambda/3} + (1 - 2\sqrt{\lambda/3}) \exp(-2\sqrt{3\lambda}\tau_1)} \right], \dots (4.6)$$

where the source function b and the scattering parameter λ may be found, in terms of the transition rates, from the method given by Giovanelli and Jefferies (1954).

The ratios N_{2^1}/N_1 and N_{2^3}/N_1 computed from (4.3) and so applying, for the singlet states, only to high opacities in $L\alpha$ and $L\beta$, are given in terms of N_e and N_4/N_1 in Table 7.

V. DISCUSSION AND APPLICATIONS TO THE SOLAR CHROMOSPHERE

Comparison of the ratios of the population of the ground state to that of the ionized state in helium and hydrogen shows that they are of the same order for a temperature T_1 in H as for a temperature $2T_1$ in He. This is as expected since the ionization potentials differ by a factor 2.

We shall now consider some aspects of solar helium emission. In the visible part of the spectrum of the undisturbed disk no trace of He absorption or emission lines is found even in the case of the strong D_3 transition, though the 2^3S-2^3P infra-red line at $\lambda 10830$ appears in absorption, indicating very different optical depths in these lines.

The optical thickness in the D_3 , 3^3D-2^3P , line may be obtained from well-known expressions for the Doppler broadened absorption coefficient and, at the line centre, is given by

$$\tau(D_3) = \int_0^\infty \frac{8.4 \times 10^{-11}}{\sqrt{T}} N_{2^3P} dz, \quad \dots\dots\dots (5.1)$$

the integration extending throughout the chromosphere. For the $\lambda 10830$ line,

$$\tau(10830) = \int_0^\infty \frac{1.4 \times 10^{-10}}{\sqrt{T}} N_{2^3S} dz. \quad \dots\dots\dots (5.2)$$

We shall consider firstly the ratio of the optical depths in D_3 and $\lambda 10830$. For simplicity we have throughout assumed the states of principal quantum number 2 to be degenerate. In fact this is not so; the 2^3S and 2^3P states have an energy difference of 1.14 eV and so the two populations are not in the ratio of their statistical weights. The relative populations are maintained primarily by radiative transitions, so that if the radiation in the atmosphere corresponds at $\lambda 10830$ to black-body radiation at a temperature T °K diluted by a factor κ it follows that

$$\frac{N_{2^3P}}{N_{2^3S}} \simeq \frac{\omega_{2^3P}}{\omega_{2^3S}} \frac{\kappa}{\exp(E_0/kT) - 1}, \quad \dots\dots\dots (5.3)$$

where $E_0 = 1.14$ eV. With $T = 6000$ °K and $\kappa = \frac{1}{2}$, the ratio of the optical depths given by (5.1) and (5.2) becomes

$$\frac{\tau(D_3)}{\tau(10830)} \simeq \frac{1}{10}. \quad \dots\dots\dots (5.4)$$

According to Mohler *et al.* (1950) $\lambda 10830$ is depressed about 5 per cent. below the local continuum, meaning that the chromospheric optical depth in this line is

~ 0.05 . From (5.4), the optical depth in D_3 should then be ~ 0.005 which is too small to allow D_3 absorption to be seen.

The magnitudes of the optical depths may be used to derive an estimate of the chromospheric temperature. In the solar chromosphere the concentration and distribution of helium is poorly known. To estimate optical depths, we use the homogeneous model defined in Section III, which should give results of roughly correct order. For $T=10^4$, 1.25×10^4 , and 1.5×10^4 °K the optical depths in D_3 and $\lambda 10830$ are then found, taking $N_{2^3S}=N_2/4$ and using (5.4), to be as shown in Table 8. This table shows that at temperatures above about 1.1×10^4 °K absorption lines should appear in $\lambda 10830$ and at about 2000 °K higher they should appear in D_3 .

TABLE 8
OPTICAL DEPTHS IN D_3 AND $\lambda 10830$ FOR THE MODEL IN SECTION III

T (10^4 °K)	$\tau (D_3)$	$\tau (10830)$
1.00	2×10^{-4}	2×10^{-3}
1.25	4×10^{-2}	4×10^{-1}
1.50	4×10^{-1}	4

While these results indicate that the mean kinetic temperature of the regions responsible for helium emission in the lower chromosphere is of the order of 1.25×10^4 °K or less, the shortcomings of the model make this estimate a rather crude one. A better indication of the temperature can be obtained from measurements of the helium emission in the chromospheric spectrum such, for example, as those of Perepelkin and Melnikov (1935) on the D_3 line. According to these authors, the emission of D_3 at a height of 1000 km above the solar limb is approximately 10^{-4} ergs $\text{cm}^{-3} \text{sec}^{-1}$. The population N_{3^3D} of the upper level of D_3 is then given by

$$Ah\nu N_{3^3D} \sim 10^{-4}, \quad \dots \dots \dots (5.5)$$

where A ($=7.2 \times 10^7 \text{sec}^{-1}$) is the spontaneous transition probability for D_3 . From (5.5) we find

$$N_{3^3D} \sim 4 \times 10^{-1} \text{cm}^{-3}. \quad \dots \dots \dots (5.6)$$

This result may be used, as follows, to infer a value of the kinetic temperature in the solar chromosphere at a height of 1000 km above the solar limb. At this height the electron concentration is, from observation, $\sim 2 \times 10^{11} \text{cm}^{-3}$. We assume that these electrons arise entirely from ionization of hydrogen and that the H:He abundance ratio is 5. From expressions, derived in paper A, for the degree of ionization of hydrogen at various temperatures, and with similar expressions for helium obtained here, the ground state population of helium, N_1 , may be found as a function of temperature. Using the triplet analogue to (4.3) we can then compute N_{3^3} and hence $N_{3^3D}=(5/9)N_{3^3}$. In this way the values of N_{3^3D} shown in Table 9 have been obtained. Results are shown for both high

and low opacity in the Lyman type continuum although the former almost certainly applies at 1000 km in the solar chromosphere if the mean kinetic temperature is of the order of 1.5×10^4 °K or less.

Comparison of the values of N_{3^3D} in Table 9 with the result (5.6) indicates that the mean temperature is not much higher than 10^4 °K. Owing to the rapid variation of N_{3^3D} with temperature, only gross changes in the data could materially affect this result.

TABLE 9
POPULATIONS OF THE 3^3D STATE

T (10^4 °K)	N_{3^3D} (cm^{-3})	
	$\tau_c \gg 1$	$\tau_c \ll 1$
1.00	4×10^{-1}	1×10^{-2}
1.25	4×10^1	1
1.50	1×10^3	2×10^1
2.00	6×10^4	6×10^2

Whether or not this temperature estimate is a unique one depends on the corresponding values of N_{3^3D} below 10^4 °K. These cannot be obtained directly from the computations here or in paper A as these do not extend below 10^4 °K. We can, however, estimate the trend as follows. At 5×10^3 °K, the various populations should be the order of magnitude given by the Saha relation. With $N_e = 2 \times 10^{11} \text{ cm}^{-3}$ and the abundance ratio used above, it is found that $N_{3^3D} \sim 5 \times 10^{-8} \text{ cm}^{-3}$ at 5×10^3 °K. It follows that, for a constant electron concentration, N_{3^3D} decreases quite rapidly below 10^4 °K, and that the observations are uniquely satisfied by a temperature close to 1.0×10^4 °K.

Perepelkin and Melnikov's observations, on which the above has been based, are of considerable interest in showing a maximum emission of D_3 at a height of 2000 km above the solar limb. It seems probable that this is due to an increase of kinetic temperature with height and the results could be analysed to determine this temperature gradient from the helium abundance. An attempt along these lines using reasonable values of the latter two parameters has been made, but indicates little beyond the fact that the mean temperature in the first 4000–5000 km of the solar chromosphere is in the range 10^4 – 1.5×10^4 °K if the helium abundance at the base is of the order of $5 \times 10^{10} \text{ cm}^{-3}$.

VI. ACKNOWLEDGMENT

The author would like to express his thanks to Dr. R. G. Giovanelli for his interest and advice.

VII. REFERENCES

- BATES, D. R., and DAMGAARD, A. (1949).—*Phil. Trans.* **242**: 101.
 BATES, D. R., FUNDAMINSKY, A., LEECH, J. W., and MASSEY, H. S. W. (1950).—*Phil. Trans.* **243**: 93.
 GIOVANELLI, R. G. (1948).—*Aust. J. Sci. Res.* **1**: 275.
 GIOVANELLI, R. G., and JEFFERIES, J. T. (1954).—*Aust. J. Phys.* **7**: 570.

- HOPFIELD, J. J. (1930).—*Astrophys. J.* **72**: 133.
- JEFFERIES, J. T. (1955).—*Mon. Not. R. Astr. Soc.* (in press).
- JEFFERIES, J. T., and GIOVANELLI, R. G. (1954).—*Aust. J. Phys.* **7**: 574.
- MENZEL, D. H. (1931).—*Lick Obs. Pubs.* **17**; Part 1.
- MIYAMOTO, S. (1951).—*Publ. Astr. Soc. Japan* **2**: 113.
- MOHLER, O. C., PIERCE, A. K., McMATH, R. R., and GOLDBERG, L. (1950).—"Photometric Atlas of the Near Infra-Red Solar Spectrum." (Univ. Michigan Press: Ann Arbor.)
- PEREPELKIN, E. J., and MELNIKOV, O. A. (1935).—*Pulkovo Obs. Bull.* **14** (5).
- REDMAN, R. O. (1942).—*Mon. Not. R. Astr. Soc.* **102**: 140.

GALACTIC SURVEY AT 400 Mc/s BETWEEN DECLINATIONS —17° AND —49°

By R. X. MCGEE,* O. B. SLEE,* and G. J. STANLEY*

[*Manuscript received February 16, 1955*]

Summary

The zone of the sky between declinations —17° and —49° was surveyed with a pencil-beam aerial. The operating frequency was 400 Mc/s and the aerial beam width 2° between half-power points.

It is shown by means of contour diagrams of equivalent sky temperature that the central Milky Way consists of a narrow background of radiation (about 5° between half-power points) on which is superimposed a number of discrete radio sources. Test measurements at two points on the galactic plane indicate random polarization of the radiation.

A most intense source in R.A. 17 hr 42 min, Dec. —28·5° (1950) is believed to represent radiation from the galactic nucleus.

The radio-frequency spectra of the Milky Way sources and other sources observed in the course of the survey are shown to be divided into two classes, indicating two different mechanisms in the production of discrete source radiation.

I. INTRODUCTION

In general, surveys of galactic radio-frequency radiation have been deficient in important respects. On the one hand, wide-beam-aerial surveys have produced rather “smoothed out” contours of radio brightness, and on the other, interferometer surveys, while giving lists of many radio sources, have failed to show how these sources are associated with the background. Thus it has become obvious that improved methods such as pencil-beam surveys of high resolution, pioneered by Hanbury Brown and Hazard, are necessary before a balanced radio picture of the sky can be produced.

The present survey was undertaken to apply the pencil-beam technique to a sample of sky of outstanding interest, that region of the Milky Way in the direction of the galactic centre.

Reber (1944, 1948), Bolton and Westfold (1950), and others have shown that the contours of radio brightness ascend to a rather broad peak close to the calculated position of the galactic centre. Bolton *et al.* (1954), using two types of interferometer and a 6° pencil beam in the Scorpio-Sagittarius region, resolved the broad radio maximum into two large extended sources. One of these which they call “source L” corresponds approximately in position to the intense discrete source discovered by Piddington and Minnett (1951) at a frequency of 1210 Mc/s and inferred to be an extended physical object in the direction of the centre of the galaxy. Mills (1952) and Bolton, Stanley, and Slee (1954) observed

* Division of Radiophysics, C.S.I.R.O., University Grounds, Sydney.

this object as a discrete source with interferometers at 100 Mc/s, but it was pointed out by Bolton *et al.* (1954) that the central Milky Way region is rather unsuitable for the use of interferometers of moderately broad primary beam owing to the confusion caused by a number of sources in such a beam at the same time. The present work clears up this confusion to a large extent and indicates considerable detail in the fine structure of the Milky Way.

The survey was carried out over the zone of the sky between declinations -17° and -49° . The operating frequency was 400 Mc/s, and the angular width of the aerial beam was 2° between half-power points. A preliminary report of the observations on the region surrounding the galactic centre, which lead to the recognition that an intense observed source may well be the galactic nucleus itself, has been presented elsewhere by McGee and Bolton (1954).

The present paper begins with a short description of the aerial and receiver used. Following is a discussion of the method of calibrating the equipment in order to establish absolute values for observed temperatures so that contours of equivalent sky temperatures could in turn be drawn. Results of the survey of the central Milky Way are then summarized, with emphasis placed on the "galactic nucleus source". The radiation in this region consists of a general background on which are superimposed several radio sources, and tests at two points on the galactic plane indicate random polarization of the radiation. Finally, a list of discrete sources is presented and the more important ones and the trends of the spectra discussed. It is shown that the spectra may be divided into two classes, some displaying the usual non-thermal type of emission and others suggesting thermal emission from a gas cloud over part of the frequency range.

II. AERIAL AND RECEIVER

(a) *The Aerial*

The reflector was a paraboloid of revolution of aperture 80 ft and focal length 40 ft set with its axis vertical. The aerial beam width was 2° at 400 Mc/s, and the beam could be directed by tilting the mast supporting the feed. The reflector was hollowed out of the sandy soil at the Dover Heights field station to a maximum depth of approximately 9 ft. The surface was then consolidated with concrete and completely covered with wire netting of $\frac{1}{2}$ -in. mesh and an additional supporting structure was added above ground level to give the full 80-ft diameter. Overall surface accuracy was maintained to within 1 in. The reflector was available for use at any desired plane of polarization of the primary feed. A photograph of the aerial is shown in Plate 1.

The primary aerial consisted of a conical dipole and plane reflector, a quarter-wavelength apart, fed by means of a shielded twin-wire transmission line. The shield and aerial were supported by a 3-in. aluminium mast guyed with nylon parachute cord, a material chosen to eliminate the possibility of spurious lobes in the aerial pattern.

The mast was pivoted and counterweighted on a 3-ft frame in the centre of the reflector in such a way that the feed could be tilted in the meridian plane by adjustment of the north and south guys. The relation between the angle of

tilt of the mast and the zenith angle of the direction of maximum sensitivity of the aerial beam was determined experimentally at a number of points by observations of the radio star Centaurus-A and of the Sun as it approached the summer solstice. The method of tilting the feed to alter the direction of the aerial beam has been fully described by Brown and Hazard (1951) and Silver and Pao (1944). The experimental ratio between the angle of tilt and the zenith angle of the beam was found to be very close to -0.92 , given in the latter reference, even up to angles of 20° . For a 2° beam it was possible to observe at angular distances up to 15° on either side of the zenith without serious distortion in aerial beam shape or appreciable loss in power gain.* Silver and Pao's relation between relative gain and angle of tilt was well verified by the solar observations.

The actual setting of the mast was done with the aid of a theodolite so that in the mechanical positioning of the beam high accuracy was obtained.

(b) *The Aerial Power Diagram*

The usual method of obtaining a diagram of power directivity of a large aerial of this type is to allow a well-established and isolated celestial point source to trace out the diagram as it passes through the beam. Since there was unfortunately no simple point source available in the zone under survey by this aerial, observations on two objects of angular width comparable with the beam width were used indirectly to confirm calculations. These were the radio source Centaurus-A, which has been found to be a combination of point source and associated spread source, and the Sun, whose dimensions at 400 Mc/s are fairly accurately known.

The primary field diagram was obtained by careful measurement. The secondary aerial power pattern was then calculated using the relation derived from a well-known formula (e.g. Slater 1942, p. 258)

$$F = 2\pi \int_0^R r f(r) J_0(\beta r) dr,$$

where F is the relative power, R is the radius of the aperture of the paraboloid, r is the radial distance, $\beta = (2\pi/\lambda)(\frac{1}{2}\pi - \theta)$, θ being measured from the zenith, and $f(r)$ represents the primary field in the aperture plane.

It will be seen that radial symmetry is inferred. Corresponding values of primary field in the E and H planes were so close to one another that a mean curve, $f(r)$, was taken as a very good approximation to the primary field.

The width of the aerial beam between half-power points was $2.1 \pm 0.1^\circ$.

Observations made on Centaurus-A with several different orientations of the feed showed that the assumption of radial symmetry was justified.

* *Note added in Proof.*—Recently the diffraction theory of coma has been applied to the aerial in a semi-quantitative way and this aberration was seen to be more serious than had been indicated by geometrical optics. It is believed that such effects as the "bulges" at declinations -21° and -45° in the 50°K contour on the left-hand side of Figure 2, the appearance of source No. 9, and the low level contours in Figure 5 may be attributed to third order coma.

(c) The Receiver

The observations were made using a Dicke type 25 c/s electronic radio-frequency switch which connected a conventional superheterodyne receiver alternately to the aerial and to a reference resistor at ambient temperature. The receiver output was applied to a synchronous detector, and the noise temperature difference was indicated on a recording ammeter. This method ensures that receiver gain fluctuations have little effect on the stability of the system.

The receiver noise factor was 11. The average receiver noise fluctuations represented an equivalent temperature of 2–3 °K at all aerial temperatures.

III. CALIBRATION OF THE EQUIPMENT

In calibrating the equipment the aim was to extract equivalent sky temperatures from the observed temperatures. Firstly, it was necessary to obtain a scale factor which would refer all observations to a common temperature scale. The method used is discussed in Section III (a).

The equivalent temperature measured at the receiver input terminals includes contributions from the ground surrounding the reflector and from the aerial feeder as well as from the sky in the aerial beam. The derivation of an expression for sky temperature T_s is given in Section III (b), and in Section III (c) there follows an estimation of the ground contribution T_g . The effects caused by loss in the feeder may be measured directly.

(a) Relative Temperature Calibration

The scale factor was obtained by daily measurements of Centaurus-A, and then all observed deflections were referred to the deflection produced by Centaurus-A.

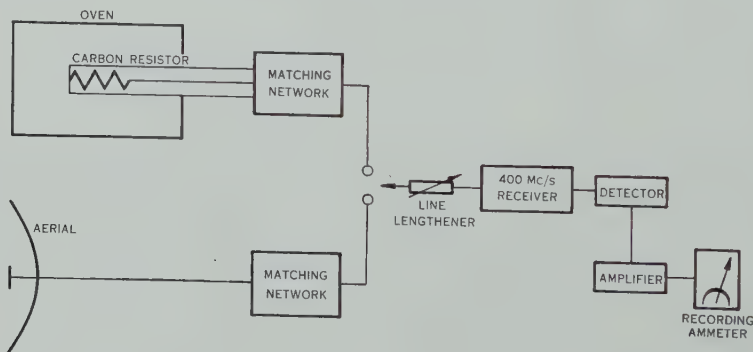


Fig. 1.—Block diagram of calibration equipment for 400 Mc/s receiving equipment.

The observed intensity of this calibrating source was then converted to an equivalent temperature by replacement of the aerial by a thermal load which could be cooled or heated over a temperature range extending from 75 °K below to 70 °K above ambient temperature. Figure 1 indicates the experimental arrangement for these measurements. The matching network of two variables

ensured that the thermal load, a carbon resistor of good high frequency characteristics mounted in a coaxial line, was very well matched to the receiver. This condition was continuously checked by the line lengthener in series with the receiver. Satisfactory matching conditions were assumed to exist when the cyclic deflections, occurring on the recording ammeter as the line lengthener was varied through a wavelength, were no longer visible above the receiver noise fluctuations. This is several orders of magnitude better than the best matching conditions indicated by the slotted line devices available on the field station.

The detector law and gain of the post-detector amplifier were checked by a carefully matched noise generator, and the recorder deflections were found to be proportional within required accuracy to power and hence to equivalent temperature over the full range of observed temperatures.

It was deduced from the measurements that the radiation from Centaurus-A at 400 Mc/s was equivalent to an aerial temperature of $74 \pm 6^\circ \text{K}$ above that of the cold sky in its immediate vicinity. This value having been determined, all observed recorder deflections could be expressed in relative temperature units.

(b) *The Equivalent Sky Temperature*

In any observation, the equivalent aerial temperature T_a is made up of the integrated beam temperature T_s , due to the radiation directed to the primary aerial by the reflector, and the unwanted ground contribution T_g resulting from thermal emission at ambient temperature, and such sky emission as may be reflected from the ground into the primary pattern.

$$T_a = T_s + T_g. \quad \dots\dots\dots (1)$$

Note that if the temperature of the sky is constant over the solid angle subtended by the secondary power diagram of the aerial, T_s may be called the equivalent sky temperature.

The quantity which can be measured directly is the equivalent temperature at the receiver input, T_r . It is different from T_a because of loss in the feeder. The aerial itself is regarded as a lossless element and of the power delivered from the aerial, $\alpha[kT_a\Delta f]$ arrives at the receiver input terminals and $(1-\alpha)[kT_a\Delta f]$ is absorbed in the feeder. The feeder is at ambient temperature, T_0 , and so contributes a power $(1-\alpha)[kT_0\Delta f]$ at the receiver input. Therefore, the power delivered to the receiver is

$$k\Delta f\{\alpha T_a + (1-\alpha)T_0\},$$

where α represents the effect of power losses in the aerial feeder, k is Boltzmann's constant, and Δf is the bandwidth. For this discussion we write the expression in terms of temperatures

$$T_r = \alpha T_a + (1-\alpha)T_0. \quad \dots\dots\dots (2)$$

Therefore, from equations (1) and (2), the required equivalent sky temperature T_s is

$$T_s = \frac{T_r - (1-\alpha)T_0}{\alpha} - T_g. \quad \dots\dots\dots (3)$$

α is obtained by measurement, T_g is estimated as in Section III (c), and so T_s may be calculated.

(c) *Estimation of the Ground Contribution, T_g*

An unwanted contribution to the observed aerial temperature occurred because the pattern of the power directivity of the primary aerial extended beyond the limits of the paraboloid aperture and over the surrounding terrain. This temperature T_g consists of a small component due to any sky radiation reflected from the ground and the main temperature T_g' , say, due to thermal emission from the ground at ambient temperature.

For the calculation of T_g' , we first assume flat ground conditions and later try to correct for the irregularities and changes in the terrain. In general,

$$T_g' = T_0 \frac{\int_0^\pi \int_0^{2\pi} A(\theta, \varphi) P(\theta, \varphi) \sin \theta \, d\theta \, d\varphi}{\int_0^\pi \int_0^{2\pi} P(\theta, \varphi) \sin \theta \, d\theta \, d\varphi}, \quad \dots \dots (4)$$

where $P(\theta, \varphi)$ is the power diagram of the primary aerial, $A(\theta, \varphi)$ is the power absorption coefficient of the surrounding medium, and θ is the angle subtended at the feed by a radius of the reflector aperture. The method used to integrate equation (4) is given in Appendix I.

In order to calculate the absorption coefficient $A(\theta, \varphi)$ it was necessary to know the dielectric constant and conductivity of the soil. These constants were kindly measured on a number of representative samples by Dr. J. S. Dryden of the Division of Electrotechnology, C.S.I.R.O. The values found were dielectric constant 4 and conductivity 1×10^{-4} mhos per metre.

Evaluation of equation (4) made $T_g' = 0.095 T_0$. Corrections were then applied to take into account departures from flat ground conditions. For example, the aerial is situated quite close to the cliff edge and approximately one-sixth of the effective spillover area extends over sea-water supplying a negligible contribution at ambient temperature and a small contribution from the reflected sky radiation. It is estimated that the equivalent temperature from reflected radiation from both sea and soil was only 0.5°K in the cases used here. The low vegetation in places was considered to cause only second order effects at this frequency.

The value of T_g obtained when the region of lowest sky temperature was observed with the aerial in the zenith position was 22°K . The errors involved in this estimate amount to probably ± 20 per cent.

Further measurements were made to check on the extent of change that could be expected in T_g as the mast was tilted, but the overall effect was well within the limits quoted above.

IV. ABSOLUTE TEMPERATURES AT SOME POSITIONS IN THE SKY

(a) *Observation of the Minimum Temperature of the Cold Sky*

The cold sky, i.e. regions away from the galactic plane and the discrete sources, exhibited such slow changes of temperature that it was not considered worth while to draw contours.

Brightness contours in previous surveys, e.g. Bolton and Westfold (1950), indicate that there is a large region of minimum sky temperature in the

approximate vicinity of the south galactic pole. Since the centre of this region sweeps through the zenith position of the aerial beam it was chosen as the first point for absolute measurement.

The equipment was set up in the same manner as shown in Figure 1. Interchanges between the aerial and the calibrating resistor at ambient temperature were made every 10 min over a 2-hr period. The deflections on the recorder chart indicating the differences between sky and ambient temperatures were constant to less than 2 per cent. over all measurements. The relative calibration by heating and cooling the resistor was then repeated. T_0 remained at 285 °K throughout and the observed value of T_r , the equivalent temperature at the receiver input, was 90 ± 6 °K.

The power loss factor α , measured for both aerial feeder and matching network by two independent methods, had a value 0.77 ± 0.04 which was in good agreement with the calculated value.

Substitution in equation (2) for T_r , α , T_0 , and T_g resulted in an equivalent sky temperature of 10 ± 5 °K (p.e.) at 400 Mc/s for the cold region whose central position was R.A. 04 hr 30 min, Dec. -33.9° .

(b) Observations along the Base Level of the Milky Way Contours

The same technique was applied at positions that gave an absolute calibration for the lowest level contour drawn in the diagrams of the radio Milky Way. The positions and temperature are given in Section V (a).

V. SURVEY OF THE MILKY WAY

The main effort in this survey was concentrated on the Milky Way in the zone of declinations from -17° to -49° . After the daily calibrating observation of Centaurus-A the aerial beam was directed towards the region of the sky under investigation. Checks were made from time to time to make sure that the position of the mast did not change due to the effects of varying weather conditions on the nylon guys; under the worst observing conditions the maximum movement of the dipole was held to 0.1° .

The Milky Way was observed on one fixed declination per day, changes in Right Ascension occurring as the rotation of the Earth swept the aerial beam across the sky. In a preliminary survey observations were made at intervals of 1° in declination. Later, in order to cover the more interesting regions in greater detail, intervals of $\frac{1}{2}^\circ$ were used. The record obtained on a particular declination was repeated until features of the variation of equivalent temperature with sidereal time were either satisfactorily reproduced or revealed as spurious and discarded.

For the analysis and fitting together of the 50 or so individual records the adopted unit of intensity was the deflection produced by Centaurus-A above a level through points on its observed temperature curve separated by 20 min in sidereal time. This level was selected since it was well above the complications in recorded pattern caused by the spread object which forms part of Centaurus-A, and thus errors which could arise in deciding a true base line for the source were

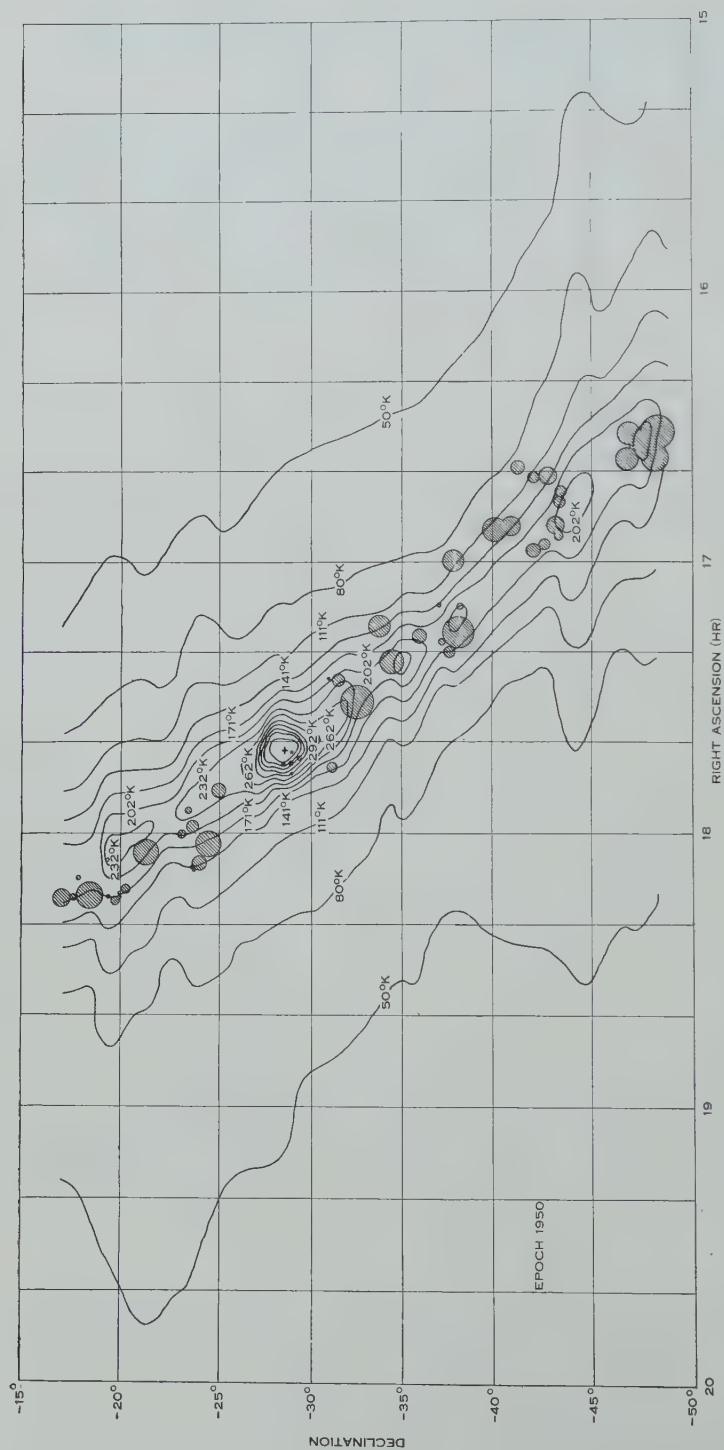


Fig. 2.—Contours of equivalent sky temperature at 400 Mc/s of central Milky Way plotted in celestial coordinates. The shaded circles represent optically observed emission regions.

eliminated. When the calibrating procedure described in Section III (a) had been carried out, it was found that one "Centaurus unit" represented an equivalent aerial temperature of 61°K .

(a) *The Milky Way Contours*

Contours of aerial temperature were plotted directly in celestial coordinates (epoch 1950) and the final diagram is presented in Figure 2. Here the contour interval is 30.5°K . The sky temperature along the first contour was checked at declinations -49° , -40° , -30° , and -20° by the absolute calibration technique referred to in Section IV (b).

For a general discussion the contours have been replotted in galactic coordinates. Precession corrections were applied to the celestial coordinates to transfer to epoch 1900 and the Lund Observatory Tables, based on the galactic pole R.A. 12 hr 40 min, Dec. $+28^\circ$ (1900), were used for the conversion. This plot appears in Figure 3, accompanied by the power diagram of the aerial.

The ability of the aerial to separate out some of the radio fine structure of the galaxy is now apparent. There are a number of intensity maxima representing discrete sources superimposed along the radio galactic plane. Particulars of these are given in Section VI (b). Studying what appear to be the contours of the background radiation, one can observe a fairly uniform increase in brightness from each side towards galactic longitude $l=328^\circ$. It is estimated that here the background radiation from the Milky Way reaches a sky temperature of about 260°K .

The diagram shows that almost all of the 400 Mc/s Milky Way in this zone is confined between $\pm 10^\circ$ in galactic latitude. The mean width between half-power points is, in fact, only 5° over the observed range of l . The centre line of the marked concentration of radio brightness runs approximately along galactic latitude -1° . A fuller discussion of the shape of the observed contours will be given in a later paper.

(b) *The Galactic Nucleus Source*

The outstanding difference from previous radio surveys is the manner in which the prominent discrete source near the galactic centre emerges from the rest of the Milky Way. This source is probably the galactic nucleus itself. Its position is very close to the generally accepted position of the galactic centre, its coordinates being $l=327.9^\circ$, $b=-1.0^\circ$, with an estimated probable error in position of $\pm 0.2^\circ$ in each case.

Statistical analysis of the counts of globular clusters, variable stars, and distant highly luminous objects such as planetary nebulae and novae, coupled with results obtained from measurements of galactic rotation and the motion of high velocity objects, place the direction of the galactic centre at 327° in galactic longitude with an uncertainty of about $\pm 1^\circ$. The optical analysis of the region is, however, difficult because parts are heavily obscured, up to 8 magnitudes in some places, by clouds associated with the Ophiucus dark nebula. They include an area containing the very positions quoted above and as yet even infra-red observations have failed to locate the galactic nucleus.

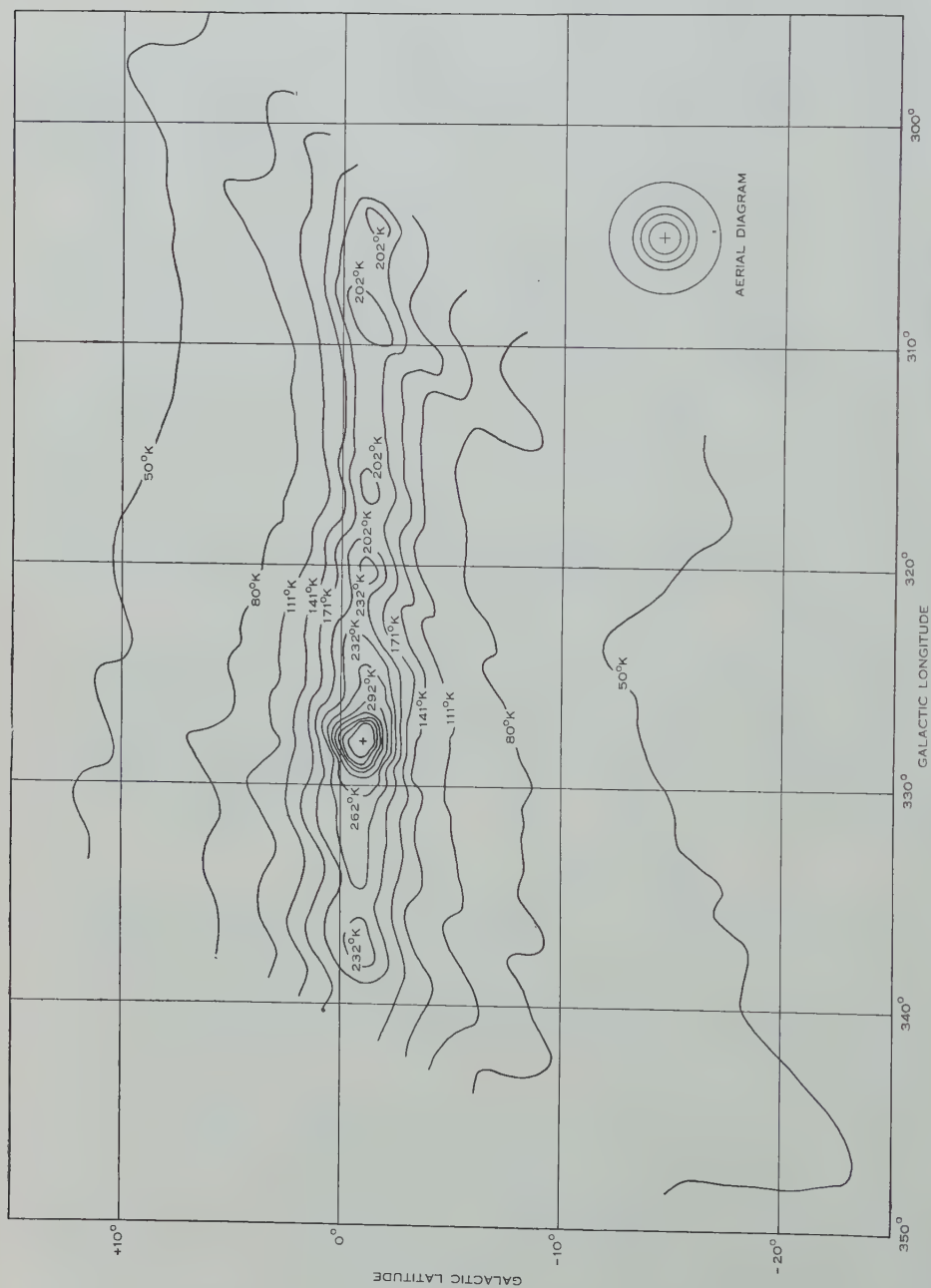


Fig. 3.—Contours of equivalent sky temperature at 400 Mc/s of central Milky Way plotted in galactic coordinates. At bottom right is a contour representation of the aerial power diagram. The circles represent the loci of the points at which received power drops to $\frac{1}{4}$, $\frac{1}{2}$, and 0 (computed) of its value for the centre of the beam.

Until recently radio estimations of the position of the centre have been subject to as much uncertainty as have optical estimations. However, since this present survey was completed, a number of other high resolution observations have been made and in these there is agreement to within ± 0.5 min in Right Ascension and $\pm 0.25^\circ$ in declination. The details are summarized in Table 1.

TABLE 1
RECENT RADIO OBSERVATIONS OF PROBABLE GALACTIC NUCLEUS SOURCE

Frequency (Mc/s)	Position 1950		Flux Density ($\text{W m}^{-2} (\text{c/s})^{-1}$)	Aerial Beam Width at Half Power	Observers
	R.A. (hr min)	Dec. (deg)			
250	17 43	-28.5	10×10^{-24}	1.2° in R.A. 17° in Dec.	Kraus, Ko, and Matt (1954)
400	17 42	-28.5	16.4×10^{-24}	2.1°	Present authors
760	17 42	-28.5	—	1.2°	Present authors (un- published data)
1420	17 43	-28.75	17×10^{-24}	0.9°	Hagen, McClain, and Hepburn (1954)
3200	17 42.5	-29.0	4.8×10^{-24}	0.4° in R.A. 0.45° in Dec.	Haddock, Mayer, and Sloanaker (1954)

Knowledge of the size would be a most important factor in identifying this source with the galactic nucleus. Dr. Baade (private communication) has pointed out that for the Andromeda nebula M31, which is probably a close counterpart of our own galaxy, the semi-stellar nucleus has a diameter of 2.5 sec of arc. He therefore expects that the nucleus of our galaxy, which is 73 times closer, would have a diameter of 3 min of arc.

Up to the present the radio data on size is conflicting. In the 400 Mc/s survey (beam width 2°) the source appears as a point source when separated from the background. The separation was effected by studying profiles taken in several directions through the point of maximum intensity. It was estimated that the general Milky Way radiation rises to an equivalent sky temperature of 260 °K and that the "nucleus source" centred on the position of the maximum produces a further equivalent temperature of 180 °K. The authors also made some exploratory observations at 760 Mc/s with an aerial beam of 1.2° between half-power points but again failed to resolve the source. Moreover, from the information published by Haddock, Mayer, and Sloanaker (1954), it is inferred that the source was effectively a point source even to their 0.4° beam at 3200 Mc/s.

The confusion in size arises at the lower frequencies and is probably due to the use of interferometers and low resolution aerials to survey a region whose structure is so complex. The source, as observed at 100 Mc/s by the interferometers of Mills (1952) and Bolton, Stanley, and Slee (1954), would appear to be small in size. But Bolton *et al.* (1954), using a 6° pencil beam and an azimuth interferometer, report that the source is 12° by 2° to fifth-power points

with a strong central concentration. Shain and Higgins (1954) infer from scintillation observations at 18.3 Mc/s that the angular size to half power is 1° or greater.

Turning now to a discussion of the intensity of radiation from the nucleus source, the flux density is plotted against frequency for the determinations listed in Table 1 together with those at 18.3 Mc/s (Shain and Higgins 1954), 100 Mc/s (Mills 1952; Bolton, Stanley, and Slee 1954; Bolton *et al.* 1954) and 1210 Mc/s (Piddington and Minnett 1951). The order of the resolution in each case is indicated by the relative size of the symbols in Figure 4, circles

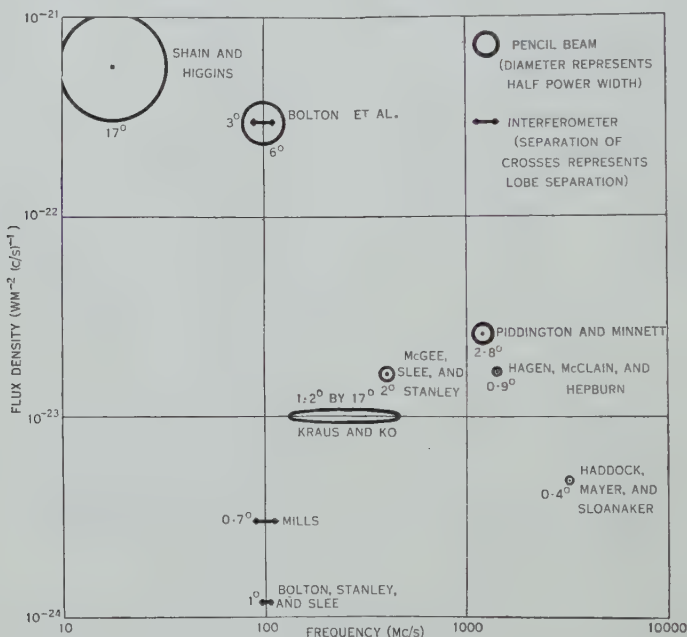


Fig. 4.—Radio-frequency radiation from the "galactic nucleus source".

being used to denote pencil-beam surveys and bars to denote interferometer surveys. A smooth spectrum curve could be drawn through the points above 250 Mc/s in frequency, especially since there are uncertainties of at least 20 per cent. in all the flux density values given. However, at 100 Mc/s, although the two interferometer points confirm one another within experimental error, the third point at this frequency, already a lower limit, has a flux density which is greater by a factor of 100. Under the circumstances no attempt is made to draw in a possible spectrum. It may be noted, though, on the one hand, if Mills and Bolton, Stanley, and Slee observed the same source as did Haddock, Mayer, and Sloanaker (unresolved with a 24' beam), there is no reason to suspect that their values of flux density would be too low, and on the other, that Piddington and Minnett (1951) published for this source a smoothly descending spectrum which is similar to the first class of spectra discussed in Section VI (c). The latter spectrum is apparently supported by the point at 18.3 Mc/s.

Because of these discrepancies in size and intensity it appears possible that the nucleus source may have a complex structure with different parts of it displaying differing spectra. High resolution surveys at frequencies near 100 Mc/s and 20 Mc/s can make important contributions towards clearing up the present confusion.

(c) *Polarization Measurements at 400 Mc/s at Two Regions in the Milky Way*

Two regions were selected for some trial polarization measurements. They were centred on Dec. -23.5° and -42.75° at the radio galactic plane. The method was the simple one of twisting the primary dipole and reflector into a number of positions removed from the original and checking for any change in the recorded temperatures in any of the positions.

On each of five successive nights the following series of observations was taken: the usual Centaurus-A transit at Dec. -42.75° for relative calibration, the transit of the galaxy at the same declination, and the transit of the galaxy at Dec. -23.5° . On the first night the equipment was set up using the normal east-west polarization of the dipole. On the second night the dipole was turned through 45° into a polarization approximately parallel to the galactic plane. On the third and fourth nights the dipole was placed in the north-south plane and in the direction perpendicular to the galactic plane respectively. The measurements were concluded with the dipole in its normal operating position as on the first night. Noise factor measurements indicated that the receiver sensitivity remained unchanged (to within 2 per cent.) over this period.

The recorded equivalent temperatures of each of the galactic regions exactly coincided on every run except No. 2 (polarization parallel to the galactic plane) when the temperature increased by 1 per cent. approximately. Thus, since reproduction of sensitivity could not be guaranteed to less than 2 per cent., it is concluded that at regions centred in R.A. 16 hr 54 min, Dec. -42.75° and R.A. 17 hr 55 min, Dec. -23.5° plane polarization of the radiation at 400 Mc/s is less than 2 per cent.

VI. DISCRETE SOURCES OBSERVED AT 400 MC/s

The discrete radio sources observed in the course of this survey are set down in Table 2. The first two columns of the table give the reference number (assigned in order of increasing Right Ascension) and constellation in which the source is located. The next three columns give respectively the position in Right Ascension and declination for epoch 1950 and the estimated flux densities at 400 Mc/s for both planes of polarization in units of $10^{-24} \text{ W m}^{-2} (\text{c/s})^{-1}$. Other authors' catalogue numbers of sources which appear to correspond are noted in the sixth column, and the final column contains remarks. The table is divided into two halves: sources 1-7 are isolated discrete sources lying outside the region covered by Figures 2 and 3, while sources 8-14 result from the analysis and composite plot of the Milky Way observations.

(a) *The Isolated Sources Nos. 1-7*

Since the survey was not designed for the searching out of discrete radio sources, there was not the same attention given to establishing their presence

TABLE 2
DISCRETE SOURCES OBSERVED AT 400 MC/s

No.	Constellation	Position 1950		Estimated Flux Density at 400 Mc/s (Both Polarizations) (10^{-24} W m $^{-2}$ (c/s) $^{-1}$)	Other Catalogue Nos.*	Remarks
		R.A. (hr min)	Dec. (deg)			
1	Fornax ..	03 20 \pm 1	-37.25 \pm 0.5	1.4	B ₁ A B ₂ 10 MO3-3 SO3-4	Evidence of being slightly spread
2	Pictor ..	05 09 \pm 1 05 16 \pm 1	-42.75 \pm 0.5 -45.0 \pm 0.5	1.5	B ₂ 22 MO5-4 SO5-4	Two maxima of approximate equal intensity; slightly spread
3	Puppis-Vela ..	08 24 \pm 2 08 35 \pm 2	-43.2 \pm 1 -45.1 \pm 1	1.5	B ₁ E B ₂ 21 MO8-4 B ₁ F SO8-4	Point source on side of spread object, 15° by 15°, of maximum equivalent temperature 138 °K above cold sky. Flux density refers to point source
4	Antlia ..	09 59 \pm 1	-28.5 \pm 0.5	0.9	B ₂ 102	One record only
5	Vela ..	10 41 \pm 1	-43.7 \pm 0.5	2.0	B ₂ 50 M10-4	Elongated in galactic longitude
6	Centaurus ..	13 22.5	-42.75 \pm 0.1	6.0	B ₁ J B ₂ 6 M13-4 H7 S13-4	Point source concentric with low intensity spread source. Flux density estimated for point source only
7	Lupus ..	15 04 \pm 1	-30.9 \pm 0.5	0.6	B ₂ 110	

Sources 8-14 superimposed on galactic plane

8	Ara	16	34 ± 1	-47.7 ± 0.5	2.3	H19	Identified with H II region by Haddock, Mayer, and Sloanaker
9	Scorpius	..	17	04 ± 1	-44.4 ± 0.5	2.6	B ₁ K B ₂ 114	Elongated in R.A. lying across galaxy
10	Scorpius	..	17	13 ± 1	-38.1 ± 0.5	3.3	M17-3 H13	Identified with H II region by Haddock, Mayer, and Sloanaker
11	Scorpius	..	17	23 ± 1	-35.0 ± 0.5	5.1	B ₂ 17 H12	Identified with H II region by Haddock, Mayer, and Sloanaker
12	Sagittarius	..	17	42 ± 1	-28.5 ± 0.2	16.4	B ₁ L B ₂ 68 M17-2B H5 S18-2	Probably galactic nucleus
13	Sagittarius	..	17	59 ± 1	-21.5 ± 0.5	2.8	M17-2A H10	Identified with H II region by Haddock, Mayer, and Sloanaker
14	Sagittarius	..	18	07 ± 1	-19.6 ± 0.5	3.6	B ₂ 58	

* B₁, Bolton *et al.* (1954); B₂, Bolton, Stanley, and Slee (1954); M, Mills (1952); H, Haddock, Mayer, and Sloanaker (1954); S, Shain and Higgins (1954).

as was given to checking the features on the central region of the Milky Way. Thus quite a number of possible source observations have had to be omitted and sources 1-7 are those whose positions correspond fairly closely to previously observed sources. Some of the more important ones are discussed further.

Source No. 2 in Pictor

The observations on this source are interesting in that they reveal two maxima of approximately equal intensity. One, located in R.A. 05 hr 09 min, Dec. $-42^{\circ}75'$, corresponds to the position of source 22 observed on the sea interferometer by Bolton, Stanley, and Slee (1954), while the other, in R.A. 05 hr 16 min, Dec. $-45^{\circ}0'$, agrees approximately in declination although not in Right Ascension with Mills's source 05-4 (1952) observed with a Michelson-type interferometer. At 400 Mc/s the source is slightly wider in angular extent than the aerial diagram.

Source No. 3 in Puppis-Vela

Unfortunately only a restricted number of observations could be made on this region and good accuracy cannot be claimed for the positions given. At 400 Mc/s an extended source covers a large area of the sky approximately bounded by 08 hr 05 min and 09 hr 10 min in Right Ascension and by -38° and -48° in declination. It lies in the $H\alpha$ emission region reported by Gum (1952), but there is no sign of correlation between the radio features and the stronger emission patches.

Superimposed on the general pattern of the observed radiation is a point source in R.A. 08 hr 24 min, Dec. $-43^{\circ}2'$. This is most likely to be Puppis-A, identified by Baade and Minkowski (1954) with an unusual nebulosity in R.A. 08 hr 20 min, Dec. $-42^{\circ}8'$. The extended source reaches a maximum equivalent temperature of 138 $^{\circ}$ K above the surrounding cold sky temperature in R.A. 08 hr 35 min, Dec. $-45^{\circ}1'$. Neither of these radio positions is in good agreement with results by observers at other frequencies.

Source No. 6 in Centaurus

This is Centaurus-A, NGC5128, which has been extensively examined since its discovery and suggested identification by Bolton, Stanley, and Slee (1949). The R.A., 13 hr 22.5 min, was taken as a reference for this survey and the Dec., $-42^{\circ}75'$, was that of the observed maximum brightness. The region was carefully observed and analysis of the records enabled the contour diagram of Figure 5 to be constructed. The contours are spaced at intervals of approximately 9 $^{\circ}$ K in equivalent aerial temperature and are plotted in celestial coordinates (epoch 1950). They have not been corrected for the effects of the aerial beam. The source appears as a spread object of approximate dimensions 5° by 3° to half power. Mills (1953) and Bolton *et al.* (1954) have referred to the fact that Centaurus-A consists of a point source and an associated spread object. In the present survey the point source is not separately distinguished.

(b) The Central Milky Way Sources Nos. 8-14

These sources are the maxima appearing along the surveyed zone of the Milky Way. The discrete source radiation merges into the general radiation

and it is difficult to decide where a separation should be made. In extracting a source from the "background" the following method was employed. The brightness contours in Figure 3 were imagined continued through the particular source under analysis according to the trend obtaining on either side, and points of equal excess brightness over them were joined to give a second system of contours, those of the "superimposed" source. With the exception of No. 9 all sources were unresolved by the 2° beam, and for flux density estimations they have been considered as point sources. All have been reported by other observers.

Observing at frequencies of 81.5 Mc/s and 210 Mc/s, Scheuer and Ryle (1953) discovered a narrow feature, 2° in angular width, superimposed on the background radiation between galactic longitudes 16° – 338° . They attribute it

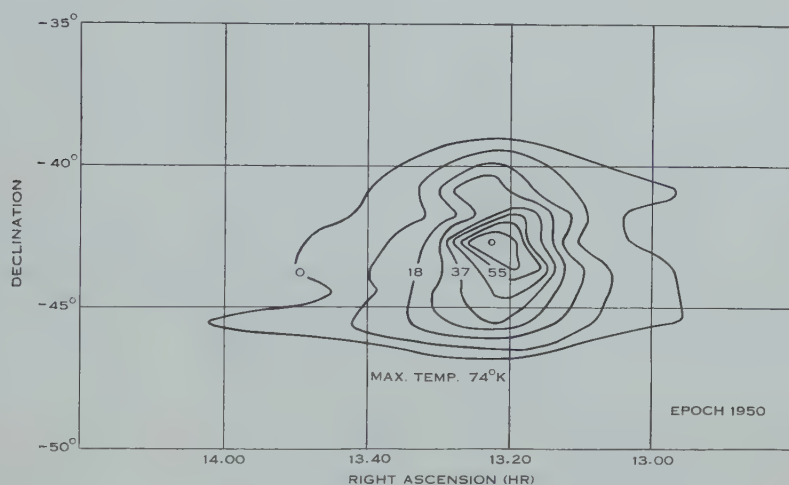


Fig. 5.—Contours of equivalent aerial temperature at 400 Mc/s of the discrete source Centaurus-A. Contour interval is 9.25°K . (Not corrected for aerial beam.)

to H II emission. In order to test whether any correlation exists between the 400 Mc/s sources Nos. 8–14 and optically observed emission regions, a comparison is made in Figure 2. The shaded circles in the figure represent emission regions published by Sharpless (1953) in the zone of declinations north of Dec. -40° , and by Gum (1953) south of Dec. -35° . It will be noticed that the H II regions are closely grouped about the "400 Mc/s galactic plane" but there is not a one-to-one correspondence with the radio sources. This may be due to insufficient aerial resolution. Haddock, Mayer, and Sloanaker (1954) with their $24'$ beam width at 3200 Mc/s have positively identified the positions of four of the sources in Table 2 with H II emission regions. However, on the present evidence concerning the radiation spectrum, to be discussed in Section VI (c), it cannot be claimed that these sources are entirely due to thermal emission. More accurate spectrum information obtained by high resolution surveys at various frequencies is desirable for deciding whether or not the origin of this radiation is H II emission.

Source No. 9 in Scorpius

Reference to the contour diagrams of Figures 2 and 3 shows that source No. 9 is an extended line of increased brightness inclined to the galactic plane. But the fact that the line is approximately coincident with the parallel of declination -44° suggests the possibility of a spurious effect, e.g. an unaccounted change in receiver sensitivity over the observing period or an unusual aerial response when the mast is in this particular setting. No explanation of this nature was established. Several observations were taken on the central declination (-44.5°) and each time a similar pattern was obtained. Also closer spaced observations on either side of this declination merely confirmed the structure as plotted.

(c) Discussion of the Radio-frequency Spectra

All of the discrete sources under discussion have been observed by other workers at one or more different frequencies (see references in Table 2), and so the opportunity has been taken to examine trends in the radio-frequency spectra. In Figure 6 observed flux density in $\text{W m}^{-2} (\text{c/s})^{-1}$ is plotted against frequency in Mc/s. Although the diagram suffers from the fact that in most cases there are too few points through which the spectrum curves are drawn it serves to illustrate the striking feature that the spectra of the 13 sources are divided into two distinct classes.

One class, characterized by a rapid decrease of flux density with frequency up to about 100 Mc/s and a subsequent gradual flattening out, is most fully represented by the spectrum for source No. 6, Centaurus-A. This curve is an extended version of the one published by Piddington and Minnett (1951). Closely following the same shape are the spectra of the well-known sources No. 1, Fornax-A; No. 2, Pictor-A; and No. 3, Puppis-A. Source No. 7 shows a similar curve over the range for which it is known.

The second class of spectra includes those of sources Nos. 4, 5, 9, 11, 13, and 14. The available data start at 100 Mc/s. The spectra exhibit an increase in flux density between 100 and 400 Mc/s followed by apparently the same type of gradual decrease towards high frequencies as in the first class. However, the increase may go past 400 Mc/s in the manner shown by the spectrum of No. 13. Sources Nos. 11 and 13 have been identified as H II regions by Haddock, Mayer, and Sloanaker (1954) while Nos. 9 and 14 may be associated with emission regions shown in Figure 2. At the low frequencies the spectra are similar to that of thermal emission from a cloud of ionized gas discussed, for example, by Piddington and Minnett (1952). Source No. 10 has a rather flat spectrum compared with either class but is identified with an H II region at 3200 Mc/s.

This division into two classes of observed spectra must correspond to an essential difference in the nature of the sources in the two groups. Two possible explanations are suggested, and in considering them one is reminded of the difficulties encountered in the nucleus source spectrum discussed in Section V (b).

One alternative is that a real difference in the spectra exists, thus giving evidence of a different mechanism in the production of radio-frequency radiation in each class of source.

GALACTIC SURVEY AT 400 MC/S



The 80-ft radio telescope.

The other, that the difference in spectra is only apparent, depends on the possibility that there has been an underestimation of flux densities at 100 Mc/s in the second class of sources. The observations at this frequency were made with interferometers of wide spacing and, if it happens that all these sources are systematically large, say, of the same order as the fringe spacing in angular

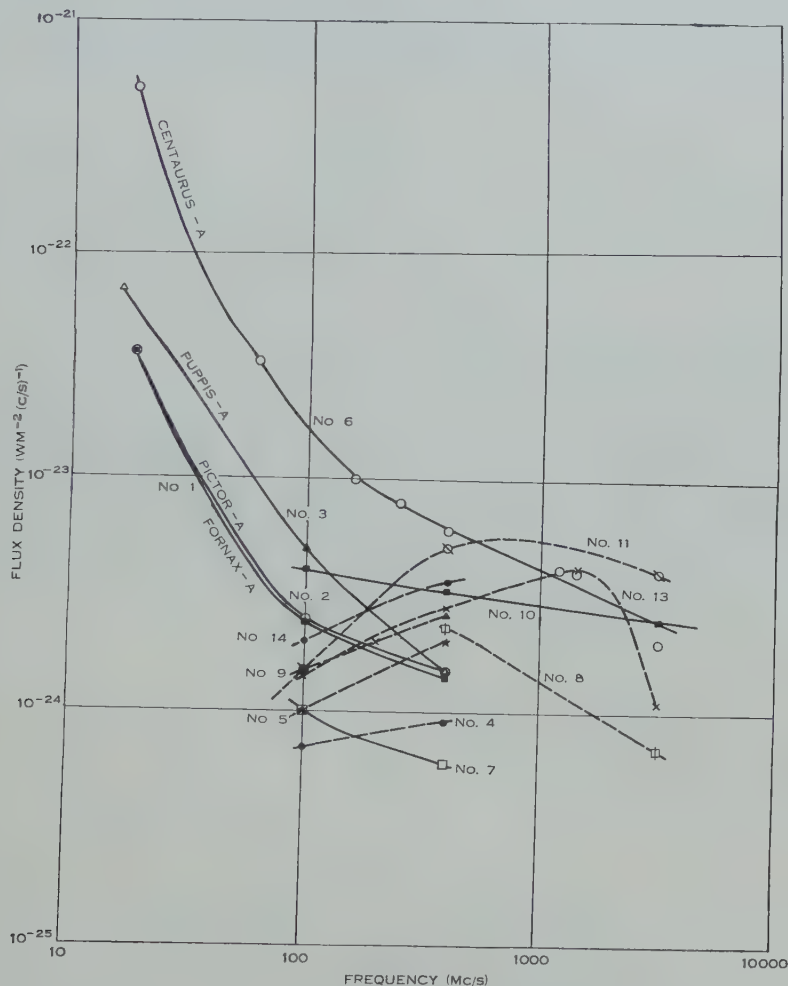


Fig. 6.—Radio-frequency spectra of 13 discrete radio sources.

size, then the interferometer observations would indicate intensities which were too low. The true flux densities may then be of such values that no division of spectra is warranted.

The question can be decided either when the true sizes of the second group have been established and all the flux densities are modified accordingly, or by observing the sources at similar resolving power to the present survey at frequencies near 100 Mc/s and lower. It is hoped that new instruments already

developed in this Laboratory, such as the Mills's "Cross" at 85 Mc/s, will soon supply the answer.

VII. ACKNOWLEDGMENTS

The authors pay tribute to Mr. J. G. Bolton of this Laboratory, who initiated the survey and was always ready to assist in the work. They wish to thank their Group Leader, Dr. J. L. Pawsey, for his advice on calibration methods and most helpful criticism of the manuscript. Finally, the discussions and criticisms by the visiting astronomer, Martha E. Stahr-Carpenter, and by other colleagues are sincerely appreciated.

VIII. REFERENCES

- BAADE, W., and MINKOWSKI, R. (1954).—*Astrophys. J.* **119**: 206–14.
 BOLTON, J. G., STANLEY, G. J., and SLEE, O. B. (1949).—*Nature* **164**: 101.
 BOLTON, J. G., STANLEY, G. J., and SLEE, O. B. (1954).—*Aust. J. Phys.* **7**: 110–29.
 BOLTON, J. G., and WESTFOLD, K. C. (1950).—*Aust. J. Sci. Res. A* **3**: 19–33.
 BOLTON, J. G., WESTFOLD, K. C., STANLEY, G. J., and SLEE, O. B. (1954).—*Aust. J. Phys.* **7**: 96–109.
 BROWN, R. H., and HAZARD, C. (1951).—*Mon. Not. R. Astr. Soc.* **111**: 357–67.
 GUM, C. S. (1952).—*Observatory* **72**: 151–4.
 GUM, C. S. (1953).—*Observatory* **73**: 123–5.
 HADDOCK, F. T., MAYER, C. H., and SLOANAKER, R. M. (1954).—*Nature* **174**: 176.
 HAGEN, J. P., MCCLAIN, E. F., and HEPBURN, NANNIELOU (1954).—*Astr. J.* **59**: 323.
 KRAUS, J. D., KO, H. C., and MATT, S. (1954).—*Astr. J.* **59**: 439–43.
 MCGEE, R. X., and BOLTON, J. G. (1954).—*Nature* **173**: 985.
 MILLS, B. Y. (1952).—*Aust. J. Sci. Res. A* **5**: 266–87.
 MILLS, B. Y. (1953).—*Aust. J. Phys.* **6**: 452–70.
 PIDDINGTON, J. H., and MINNETT, H. C. (1951).—*Aust. J. Sci. Res. A* **4**: 459–75.
 PIDDINGTON, J. H., and MINNETT, H. C. (1952).—*Aust. J. Sci. Res. A* **5**: 17–31.
 REBER, G. (1944).—*Astrophys. J.* **100**: 279.
 REBER, G. (1948).—*Proc. Inst. Radio Engrs., N.Y.* **36**: 1215–8.
 SCHEUER, P. A. G., and RYLE, M. (1953).—*Mon. Not. R. Astr. Soc.* **113**: 3–17.
 SHAIN, C. A., and HIGGINS, C. S. (1954).—*Aust. J. Phys.* **7**: 130–49.
 SHARPLESS, S. (1953).—*Astrophys. J.* **118**: 362–9.
 SILVER, S., and PAO, C. S. (1944).—M.I.T. Radiation Lab. Rep. 479.
 SLATER, J. C. (1942).—"Microwave Transmission." (McGraw-Hill: New York.)

APPENDIX I

Calculation of Aerial Temperature due to Thermal Emission from the Ground Surrounding the Paraboloidal Reflector

In Section III (c) equation (4) states that the temperature contribution T_g' at the feed and the ambient temperature T_0 are in the ratio of the power received by the feed from the surrounding medium other than the reflector itself to the power that would be received were the feed completely enclosed in a black body space at ambient temperature.

Equation (4) may be rewritten in the form

$$T_g' = T_0 \frac{\int_0^\pi \int_0^{2\pi} \{A_{||}(\theta, \varphi) P_{||}(\theta, \varphi) + A_{\perp}(\theta, \varphi) P_{\perp}(\theta, \varphi)\} \sin \theta \, d\theta \, d\varphi}{\int_0^\pi \int_0^{2\pi} P(\theta, \varphi) \sin \theta \, d\theta \, d\varphi}, \quad \dots (5)$$

in which $A(\theta, \varphi)$ and $P(\theta, \varphi)$ have been split up into components parallel (subscript \parallel) and perpendicular (subscript \perp) to the plane of incidence.

The observed power directivity pattern of the feed could be expressed analytically to a very close approximation by

$$P(\theta, \varphi) = (1 - \sin^2 \theta \cos^2 \varphi)^3 \sin^2 (\tfrac{1}{2}\pi \cos \theta). \quad \dots\dots\dots (6)$$

Then

$$P_{\parallel}(\theta, \varphi) = \sin^2 (\tfrac{1}{2}\pi \cos \theta) \{ (1 - \sin^2 \theta \cos^2 \varphi)^3 - \sin^2 \varphi \}, \quad \dots\dots (7)$$

$$P_{\perp}(\theta, \varphi) = \sin^2 (\tfrac{1}{2}\pi \cos \theta) \sin^2 \varphi.$$

Equation (5) now becomes

$$T_g' = T_0 \frac{\int_0^{2\pi} \int_0^{\frac{1}{2}\pi} \sin^2 (\tfrac{1}{2}\pi \cos \theta) \sin \theta (A_{\parallel} + A_{\perp}) \{ (1 - \sin^2 \theta \cos^2 \varphi)^3 - \sin^2 \varphi \} d\theta d\varphi}{\int_0^{2\pi} \int_0^{\frac{1}{2}\pi} \sin^2 (\tfrac{1}{2}\pi \cos \theta) \sin \theta (1 - \sin^2 \theta \cos^2 \varphi)^3 d\theta d\varphi} \quad \dots\dots\dots (8)$$

The denominator of equation (8) was evaluated by analytical integration and for the numerator integration was performed analytically with respect to φ and numerically with respect to θ .

Absorption coefficients A_{\parallel} and A_{\perp} were calculated from the usual Fresnel equations using the measured values of soil conductivity and dielectric constant.

Although integration is carried out over the full hemisphere bounded by the plane reflector, the power directivity term ensures that outside the ground over which the aerial pattern extends there is a negligible contribution to T_g' .

THE OBSERVATION AND INTERPRETATION OF RADIO EMISSION FROM SOME BRIGHT GALAXIES

By B. Y. MILLS*

[*Manuscript received April 12, 1955*]

Summary

Preliminary attempts to observe 13 bright southern galaxies are described. Of these 10 were detected, including the Magellanic Clouds. The latter were studied in detail. Supplementary measurements on the Milky Way near the galactic centre were made and also some unsuccessful attempts to observe two globular clusters. Taking these results in conjunction with some observations of Brown and Hazard it is possible to derive a consistent picture of the radio emitting properties of "normal" galaxies. It is found that radio emission probably occurs with two markedly different distributions, one displaying characteristics of a type I population, the other having a very extensive, roughly spherical, spatial distribution quite unlike any known stellar population. Both distributions have a non-thermal spectrum, and the relative contributions of each appear to vary between different galaxy types. In general terms, although not altogether in detail, these results support the conclusions reached by Shklovskii concerning the distribution of radio emission in the Galaxy; also his interpretation of the spherical distribution as the result of relativistic electrons radiating in weak magnetic fields is examined, and found, with some modifications, to afford a plausible explanation of the observed differences between the radiating efficiencies of different types of galaxies.

I. INTRODUCTION

A large radio telescope intended primarily for the study of cosmic radio-frequency radiation has recently been put into operation near Sydney. The telescope, which consists of a cruciform arrangement of two arrays of dipoles, operates at a wavelength of 3.5 m, has a beam width between half-power points of 50 min of arc and a sensitivity which, under ideal conditions, approaches $10^{-26} \text{ W m}^{-2} (\text{c/s})^{-1}$; the principle of operation has been described by Mills and Little (1953). It is arranged as a transit instrument and altered in declination by phasing the elements of the arrays. As a consequence, at zenith angles greater than about 45° , the sensitivity and resolution begin to deteriorate rapidly. The sky coverage for useful observations from Sydney (latitude 34°S.) therefore lies between declinations of about $+10^\circ$ and -80° in most circumstances. A detailed account of the equipment is being prepared for publication by Mills, Little, and Sheridan.

This radio telescope, being a pencil beam instrument, may be used to study both the discrete sources and the general background radiation, and eventually

* Division of Radiophysics, C.S.I.R.O., University Grounds, Sydney.

a detailed survey of the distribution of cosmic noise in the southern sky will be undertaken. Before commencing such a survey, however, it has been found desirable to spend some time testing and calibrating the equipment, while at the same time observing special objects of astronomical interest. This paper describes the results of observing some bright galaxies, including the Magellanic Clouds, and some regions of the Milky Way.

For the purpose of studying their radio emission, "normal" galaxies may be divided into three classes: (i) comparatively distant and faint galaxies in the magnitude range from 7 to 11 for which the only observation possible with this instrument is their total radio emission (or lack of it); (ii) close galaxies such as the Magellanic Clouds and M31 (the latter unfortunately being outside the coverage of the aerial) in which the distribution of radio emission over the galaxy may be mapped and compared with that of various stellar populations; (iii) the Milky Way which is unique in the wealth of detail which may be obtained, but in which the edge-on view leads to difficulties of interpretation requiring the use of supplementary information derived from studies of the first two types.

Attempts have been made to observe 11 galaxies of the first type of which 8 were probably detected; also both Magellanic Clouds have been observed and their radio and optical properties compared. In addition a few galactic sections have been obtained in the region of the galactic centre. Taken together these observations appear to provide a consistent picture of the distribution of radio emission throughout a galaxy and of its dependence on the constituent stellar populations. In general the observations tend to support the ideas of Shklovskii (1952, 1953) in which the galactic emission is considered to be distributed in two subsystems, one showing a discoidal distribution highly concentrated towards the galactic plane, and the other a very dispersed and approximately spherical distribution concentric with the galactic centre. However, contrary to Shklovskii's interpretation, it is found that the contribution of thermal emission from ionized hydrogen to the discoidal distribution is likely to be small.

II. RADIATION FROM THE MAGELLANIC CLOUDS

Study of the radio emission of the Clouds is difficult because both have a relatively low surface brightness compared with irregular variations in background brightness in their neighbourhood. The Small Cloud presents the greatest difficulty, particularly in its outer regions, but even in the Large Cloud the distribution of emission can be determined with confidence only near its centre. A further difficulty arises because a radio telescope of this type is relatively insensitive when observing extended distributions; consequently, very high stability is required from the equipment, and random drifts must be kept below the level of thermal noise fluctuations for the period of the observations, that is, about 7 hr. The electronic portion of the equipment meets the stability requirement, but it is not yet known how the aerial itself is affected by changes in temperature or humidity. Fortunately, the main conclusions do not depend on an exact knowledge of the detailed structure in the outer parts of the Clouds.

In Figure 1 is shown a facsimile of a record taken on the Large Magellanic Cloud at a declination of $-69^{\circ}26'$. Recordings of this type were made at declination intervals of one beam width between declinations of $-62\frac{1}{2}^{\circ}$ and $-75\frac{1}{2}^{\circ}$, together with "scanning" records over the central regions of both Clouds. In the scanning observations the aerial beam is electrically scanned back and forth over five declinations slightly less than half a beam width apart, and the section at each declination is separately recorded. It has been shown by Bracewell and Roberts (1954) that no information is lost by making measurements at intervals of less than a certain critical interval which is approximately half the beam width. Outside the central regions where scanning was not employed the critical interval was exceeded so that fine detail in the distribution may have been overlooked.

To construct a contour map from the series of sections so obtained it is desirable to obtain absolute levels of brightness. However, the calibration of the equipment was not sufficiently accurate to enable an undistorted map to be constructed, since small variations between different declinations and on different days were of the same order as the excess brightness due to the outer regions of the Clouds. Therefore a method of matching the temperatures at a particular Right Ascension has been adopted. A region from 23 hr 00 min to 23 hr 40 min has been chosen because observations with other equipment suggest that over the declination range of interest the galactic radio isophotes lie roughly parallel to an hour circle. As a first step all temperatures were equalized over the chosen period after excluding any weak discrete sources which were present. The resulting isophotes suggested strongly that north of -68° the temperatures were all too high and consequently all temperatures north of this declination were reduced by 100°K and the isophotes redrawn. The result is shown in Figure 2 (a). The contour interval is 125°K and the temperatures are all measured above a base temperature T . The value of T is estimated as approximately 700°K .*

The optical centres for the Clouds are, for the Large Cloud, 05 hr 24 min and -69.8° , and for the Small Cloud, 00 hr 51 min and -73.1° , so that in Figure 2 (a) the excess radiation near the centres of the Clouds may be readily seen. However, the isophotes also include irregularities in the background radiation and discrete radio sources so that it is difficult to determine the outer boundaries of the Clouds or their integrated emission. In Figure 2 (b) are shown the outer boundaries derived from a careful examination of the original records. The dark areas include radiation connected to the main systems by strongly closed contour lines and therefore almost certainly belonging to the Clouds.

* It has been shown that with a radio telescope of this type the absolute temperature is obtained very closely by the addition of two temperatures, one corresponding to that of the pencil beam and showing the fine detail and the other corresponding to the absolute temperature of the arrays (Mills and Little 1953). This method was used for estimating the temperature T , but variations of the array temperatures at other Right Ascensions were ignored since they have only a very small gradient. Because of this, errors of up to 100°K may be produced at any one point, but the shape of the contours would not be significantly affected.

The lighter areas are regions of excess radiation, often quite strong, but in which the connexion with the Clouds is not so obvious; they represent the extreme limits to which Cloud radiation may be detectable with the present equipment. The dotted lines are the limits from which the 21 cm line radiation from neutral hydrogen has been detected by Kerr, Hindman, and Robinson (1954). It will be noted that the bright region south of -75° stretching from one Cloud to the other shown in Figure 2 (a) has been omitted in Figure 2 (b). There is no evidence to show that it is associated with the Clouds, and from the distribution of brightness along it the alternative explanation that it represents an irregularity in the galactic radiation is more probable.

The agreement of the outer boundaries of both the 21 cm line radiation and the 3.5 m radiation over extensive regions is striking and suggests, *inter alia*, that the excess 3.5 m emission north following the main body of the Small Cloud is actually associated with it, something which is quite impossible to decide from the records themselves. The "link" between the Clouds at a

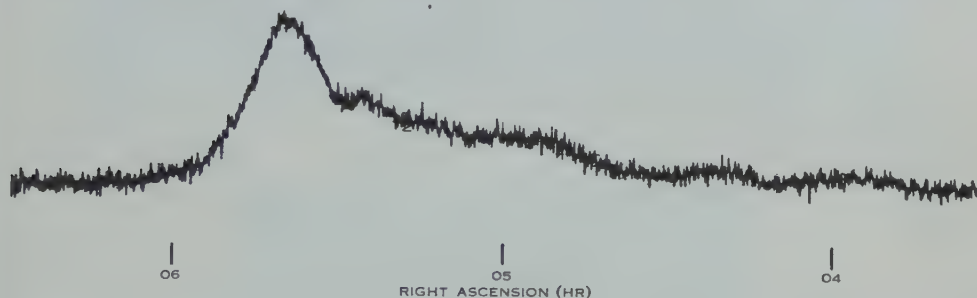


Fig. 1.—A typical record of the Large Magellanic Cloud at a declination of $69^\circ 26' \text{ S}$.

declination of -72° may not be real, as such an effect could result from a small error in the original matching of the profiles or from long-period drifts in the equipment. However, there is no internal evidence in the observations to throw doubt on its existence, so it has been retained in the diagram as a possible feature of the systems. The link is some 2° to the north of a possible link suggested by the optical evidence.

A photograph of the Clouds is shown in Plate 1, Figure 1. There appears to be little resemblance between the photographic appearance and the distribution of radio emission, because of differences in the shapes of the brighter features and the apparently greater extent of the radio distributions. However, de Vaucouleurs (1954), by means of long exposure photographs and star counts, has traced the outer regions of the Large Cloud over very wide areas, in places exceeding the radio boundaries. He finds no optical extension of the Small Cloud in a north-easterly direction, corresponding to the possible extension of radio emission.

It is interesting to compare the brighter features of the three distributions in 3.5 m continuous spectrum radiation, in 21 cm line radiation, and in visible radiation. A striking circumstance is the complete absence of the bright visible axial bar of the Large Cloud at a wavelength of 3.5 m, although it is slightly

visible in the 21 cm line radiation. A similar discrepancy in the Small Cloud, although present, is not so obvious. The greater tilt of this system makes it more difficult to observe such structural features. In both Clouds the region of maximum radio emission is displaced towards the greatest concentration of

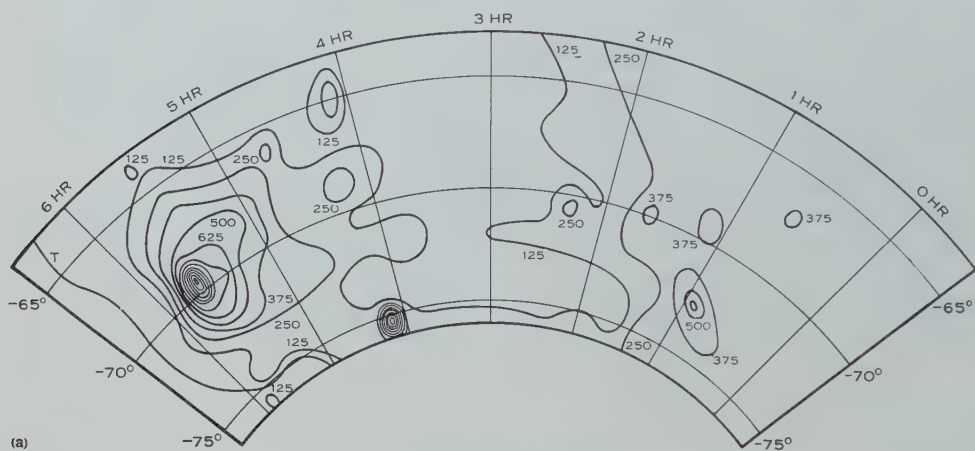


Fig. 2 (a).—Contour maps of the 3.5 m isophotes in the vicinity of the Magellanic Clouds; the contour interval is 125 °K.

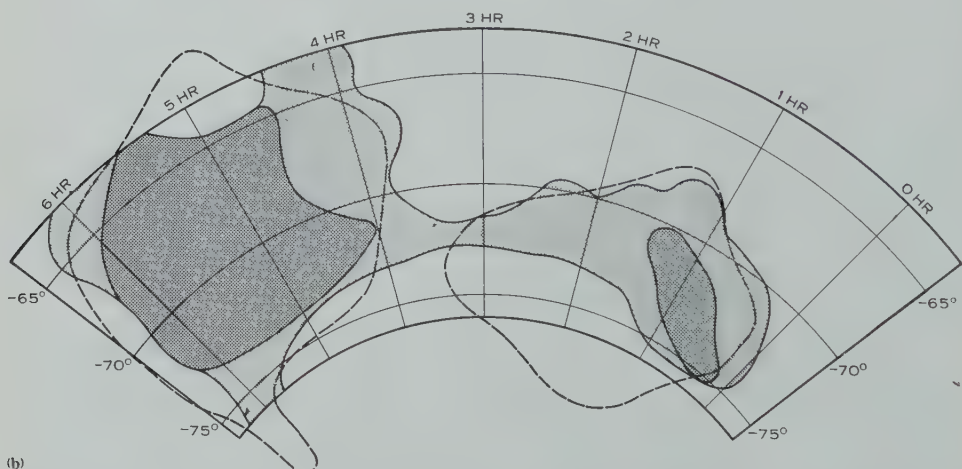


Fig. 2 (b).—The boundaries of the Clouds at 3.5 m wavelength compared with the hydrogen line emission boundaries. The dark areas represent definite regions of 3.5 m radiation and the lighter areas possible regions of the radiation. The dotted lines outline the neutral hydrogen boundaries.

bright stars ($m < 14$) which, according to de Vaucouleurs (1955*a*, 1955*b*), also show no concentration in the axial bars. However, the “wing” of the Small Cloud, although very rich in such stars, shows no trace of 3.5 m radiation. The positions of maximum brightness of the two types of radio emission agree

RADIO EMISSION FROM SOME BRIGHT GALAXIES

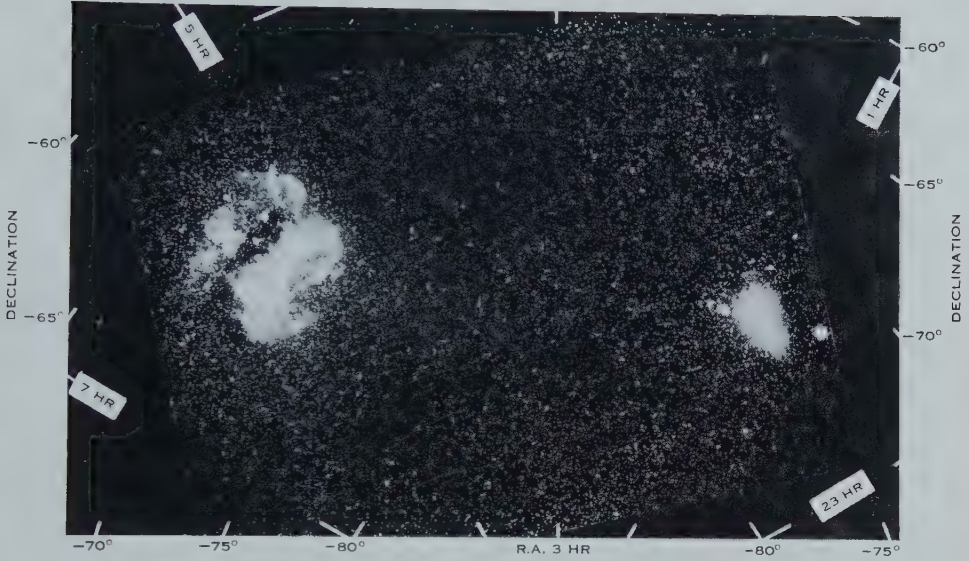


Fig. 1.—A photograph of both Magellanic Clouds.

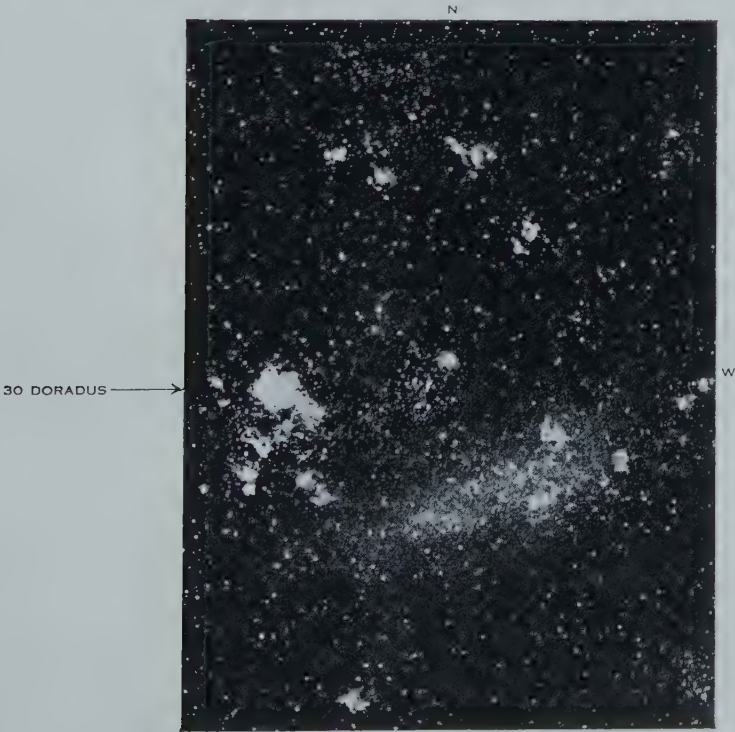


Fig. 2.—A photograph of the Large Magellanic Cloud in $H\alpha$ light.

closely, being for the Large Cloud, 05 hr 40 min, -69.6° at 3.5 m* and 05 hr 38 min, -70° for the hydrogen line emission and, for the Small Cloud, 00 hr 56 min, -72.9° at 3.5 m and 01 hr 00 min, -73.1° for the hydrogen line emission; the agreement is within the experimental error. The general distributions of the radio emissions are also similar except for the discrepancies in the axial bar of the Large Cloud and the lack of 3.5 m radiation south following the main body of the Small Cloud where there is an extensive region of neutral hydrogen.

TABLE 1
CENTRES OF THE LARGE MAGELLANIC CLOUD DISTRIBUTIONS

Centre	R.A.		Dec. (deg)
	(hr	min)	
Centroid of neutral hydrogen	05	35	-68.5
Centroid of bright stars	05	33	-67.7
Optical centre (brightest part)	05	24	-69.8
Centre of rotation (radio data)	05	20	-68.8
Centroid of 3.5 m radiation	05	17	-69.3

Comparisons may be extended further by considering the centroids of the various distributions (Table 1) and their average radial decrements. The centroid of the 3.5 m distribution in the Large Cloud has been calculated by making reasonable estimates of the percentage of radiation originating in the outer regions and is probably accurate to about half a degree. Ignorance of the distribution of weak radiation far from the centre is the chief cause of

TABLE 2
CENTRES OF THE SMALL MAGELLANIC CLOUD DISTRIBUTIONS

Centre	R.A.		Dec. (deg)
	(hr	min)	
Centroid of neutral hydrogen	01	20	-72.5
Centroid of bright stars	01	00	-72.5
Optical centre (brightest part)	00	51	-73.1
Centre of rotation (radio data)	01	10	-73.25
Centroid of 3.5 m radiation (bright regions only)	00	54	-72.7

uncertainty. Other data have been taken from de Vaucouleurs (1955*a*) and Kerr and de Vaucouleurs (1955).

All the centres differ in position, but in view of the uncertainties involved no definite conclusions can be drawn.

A similar comparison may be made for the Small Cloud (Table 2) but the uncertainty in the centroid of the 3.5 m radiation is even greater. The value

* The prominent maximum in the distribution of 3.5 m radiation near this position is thought to be due to the nebula 30 Doradus, as explained later. The maximum in the general emission remains, however, after the effect of the nebula is subtracted.

given is that of the dark area of Figure 2 (*b*). If the suspected outer component is included the centroid would be moved in a north-following direction.

Again the centres differ and the uncertainties are such that it is difficult to come to any conclusions. It does seem, however, that the main mass of the system is east of the brightest regions so that again the indications are that the 3.5 m radiation could extend in the manner suggested.

Following de Vaucouleurs (1955*a*), the distributions in the Large Cloud may be further compared by plotting average values at given angular separations from some assumed centre, in this case taken as the optical centre. It does not seem worth while carrying out this procedure for the Small Cloud, however, because of its marked elongation and asymmetry. The data for the Large Cloud are presented in Figure 3 where the number density of bright stars, the

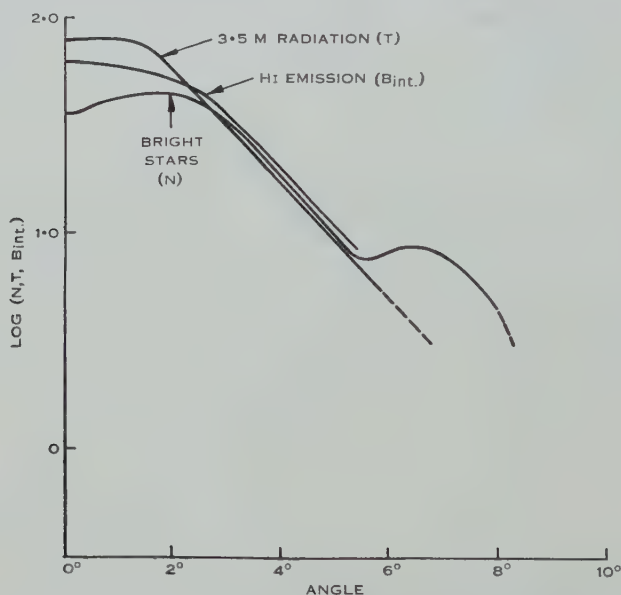


Fig. 3.—A comparison of the average radial distributions of 3.5 m radiation, 21 cm hydrogen line radiation, and bright stars in the Large Magellanic Cloud.

H line integrated brightness, and the 3.5 m temperatures are plotted as logarithmic functions of the radius, with arbitrary zeros. The general similarity is obvious, particularly for radii between 2° and 6° where the logarithmic decrements are identical. Differences towards the centre reflect the detailed differences already discussed.

Although neither these bright stars nor the neutral hydrogen show a one-to-one correspondence with the distribution of 3.5 m radiation, all the distributions have a considerable similarity, which strongly suggests that the radiation is associated with the presence of interstellar matter and of bright stars, that is, that it has a Population I type of distribution. The possibility that a large proportion of the radiation might be due to thermal emission of

ionized hydrogen therefore demands investigation. This is most readily checked in the Large Cloud. An inspection of Figure 2 (*a*) reveals a marked concentration of radiation at a position centred on R.A. 05 hr 40 min and Dec. -69.3° . The source is not resolved in an east-west direction but is slightly extended at right angles.* The excess radiation most likely originates in the giant emission nebula 30 Doradus at a position of 05 hr 39.9 min, -69.1° and the neighbouring emission regions to the south. With this interpretation the flux density of 30 Doradus is about $3 \times 10^{-25} \text{ W m}^{-2} (\text{c/s})^{-1}$ and the average of the other emission regions is about $2.5 \times 10^{-25} \text{ W m}^{-2} (\text{c/s})^{-1}$. These values seem quite consistent with thermal emission, since, if 30 Doradus is assumed to be optically thick at a wavelength of 3.5 m, of angular size 15 min, and of temperature 10,000 °K, the flux density would be $3.4 \times 10^{-25} \text{ W m}^{-2} (\text{c/s})^{-1}$.

This result may be used to show that the greater part of the radiation from the large Cloud probably has a non-thermal origin. Thus it is shown later that the integrated emission from the whole Cloud is approximately $2 \times 10^{-23} \text{ W m}^{-2} (\text{c/s})^{-1}$ so that, if this also were due to thermal emission, the mass of ionized hydrogen in 30 Doradus would be very small compared with the total mass of ionized hydrogen in the Cloud. This seems very improbable; see, for instance, a photograph of the Large Cloud taken in $\text{H}\alpha$ light by C. Gum of Mt. Stromlo Observatory (Plate 1, Fig. 2) where the nebula 30 Doradus dominates the whole picture. Recently this conclusion has received confirmation from some unpublished results of J. H. Piddington and G. Trent of the Radio-physics Laboratory who, observing with an aerial of 3° beam width at a wavelength of 50 cm, find that the only detectable excess radiation in the region is a source of small angular size close to 30 Doradus and of flux density very roughly $1.5 \times 10^{-24} \text{ W m}^{-2} (\text{c/s})^{-1}$. Since the flux density at a wavelength of 50 cm would be equal to or greater than that at 3.5 m if the emission had thermal origin, the conclusion that at 3.5 m thermal gas emission is only an insignificant proportion of the total appears inescapable. The radiation from the Small Cloud is too weak to apply similar tests, but there seems to be no reason for assuming that the results would be any different. The probable extensions of the radio emission of the Cloud in regions where no emission nebulae are observed (see, for instance, Henize and Miller 1951) support this conclusion.

A further check on the relation of radio emission to population types is afforded by the presence near the Small Cloud of two bright globular clusters, NGC 362 ($m_p = 6.0$) and 47 Tucanae ($m_p = 3.0$). Neither gives any detectable radio emission although, if the ratio of radio to optical emission in 47 Tucanae were the same as in late-type spirals, a signal nearly 500 times the minimum detectable level would be expected. It therefore appears safe to conclude that the bright type II stars found in globular clusters do not contribute significantly to the total energy of cosmic radio waves.

* At the large zenith angles at which these observations were made the aerial beam becomes elongated in a north-south direction, but the apparent extent of the source is greater than that of the aerial beam.

Summing up the results of these comparisons it appears that at a wavelength of 3.5 m the emission from the Magellanic Clouds originates principally in a non-thermal process and has a distribution which is closely related to the interstellar gas and bright stars. The bright type II stars of globular clusters are not associated with radio emission, but whether the fainter type II stars found in elliptical galaxies and the disks of spirals are so associated needs further investigation, including comparisons of the Clouds, which appear to lack such stars, with other galaxies.

III. THE INTEGRATED EMISSION OF THE CLOUDS

The total emission from each of the Clouds can be obtained by direct integration of the contours, the flux density being given by the relation

$$S = \frac{2k}{\lambda^2} \int T d\Omega.$$

The value obtained is naturally very dependent on the distribution in the outer regions where, although the temperature is low, the solid angle is very great. Values obtained assuming an extent given by (a) the minimum shown by the dark area in Figure 2 (b) and (b) the possible extent indicated by the light area in the same figure are as follows :

LMC	$S_{(a)} = 1.7 \times 10^{-23} \text{ W m}^{-2} (\text{c/s})^{-1}$
	$S_{(b)} = 2.2 \times 10^{-23} \text{ W m}^{-2} (\text{c/s})^{-1}$
SMC	$S_{(a)} = 1.15 \times 10^{-24} \text{ W m}^{-2} (\text{c/s})^{-1}$
	$S_{(b)} = 3.7 \times 10^{-24} \text{ W m}^{-2} (\text{c/s})^{-1}$

The flux density of the Large Cloud may be estimated in another way from the radial decrement curve of Figure 3. This method has been used by de Vaucouleurs (1955*a*, 1955*b*) and by Kerr and de Vaucouleurs (1955) in estimating the total stellar populations and the mass of hydrogen. The curve of $\log T$ plotted against radial angle is extrapolated to infinity as a straight line of uniform slope, and the integral performed. This integral will always converge, and in the present case leads to a value of the total emission of $1.95 \times 10^{-23} \text{ W m}^{-2} (\text{c/s})^{-1}$ in very good agreement with the results above. The value is, however, rather lower than $S_{(b)}$, indicating that not all the excess radiation in the vicinity is associated with the Cloud.

For comparisons with other galaxies it is convenient to assign definite values to the flux densities which are estimated to be as follows :

LMC	$S = 2 \times 10^{-23} \text{ W m}^{-2} (\text{c/s})^{-1} \pm 15 \%$
SMC	$S = 3 \times 10^{-24} \text{ W m}^{-2} (\text{c/s})^{-1} \pm 30 \%$

The flux density of the Large Cloud is hardly affected by the uncertainties in the outer boundaries, and the main allowance in the estimated probable error is for calibration errors in the equipment. The relative importance of the errors is reversed in the Small Cloud, and here the flux density is taken nearer the higher value as it is considered probable that most of the excess radiation within the confines of the Cloud, as defined by the hydrogen boundaries, is

associated with it. It will be shown later that this assumption is consistent with what is known of the distribution of radiation in other galaxies.*

To compare the flux densities of the Clouds with those of other galaxies it is convenient to express them in the magnitude scale of Brown and Hazard (1952), applying a correction for the difference in wavelength so that the northern and southern observations may be directly compared.

The uncorrected values are as follows :

$$\begin{aligned} \text{LMC} \quad m_{R(3.5 \text{ m})} &= -53.4 - 2.5 \log_{10} S_{(3.5 \text{ m})} \\ &= 3.35 \pm 0.15 \end{aligned}$$

$$\text{SMC} \quad m_{R(3.5 \text{ m})} = 5.4 \pm 0.35.$$

To convert to the wavelength of 1.9 m (158 Mc/s) used by Brown and Hazard it is assumed that the spectrum is of the form $S \propto \lambda^{0.7 \pm 0.1}$, which gives a difference in magnitude of 0.47 ± 0.06 . We therefore have for the radio magnitudes at 1.9 m :

$$\begin{aligned} \text{LMC} \quad m_R &= 3.8 \pm 0.2 \\ \text{SMC} \quad m_R &= 5.9 \pm 0.4. \end{aligned}$$

IV. OBSERVATIONS OF OTHER BRIGHT GALAXIES

Attempts have been made to observe eleven other bright galaxies in the southern sky. Of these, eight were probably detected and three could not be observed. Several of the galaxies were barely above the sensitivity limit of the equipment so that an accurate position measurement and certain identification could not be expected. Others are in regions of radio source concentrations, again leading to difficulties of identification and of flux density measurements.

The aerial may be directed to a series of standard declinations at approximately 20' intervals. Observations were therefore made by directing the beam to the declination nearest to that of the galaxy and observing its meridian transit. If no signal could be detected after two such transits the galaxy was noted as undetectable and an estimate made of the upper limit of its flux density. If a source was observed at the correct Right Ascension (within 0.2 min for the stronger sources and 0.5 min for the weaker ones) checking observations were made at declinations on either side. The result was a declination accuracy of only about half a degree, but it was considered that this was adequate to ensure generally that no spurious identifications were made. Intensity measurements

* The values above may be compared with some early measurements obtained with the experimental version of the aerial, of 8° beam width (Mills and Little 1953), and quoted in an U.R.S.I. report (1954). They were as follows :

$$\begin{array}{ll} \text{LMC} & S = 3.1 \times 10^{-23} \\ \text{SMC} & S = 9.0 \times 10^{-24} \end{array} \left. \vphantom{\begin{array}{l} \text{LMC} \\ \text{SMC} \end{array}} \right\} \lambda = 3.1 \text{ m}$$

It seems fairly clear that because of the low aerial resolution all the excess radiation in the vicinity was then included in both Clouds.

are accurate to about 10 or 20 per cent. and are expressed in the same magnitude scale as used for the Clouds. The results of observations of these galaxies, the Clouds, and the previously mentioned globular clusters are summarized in Table 3. The individual observations are discussed in more detail below.

The Large Magellanic Cloud

The radio observations have been discussed already. The photographic magnitude is an estimate of de Vaucouleurs (Buscombe, Gascoigne, and de Vaucouleurs 1954) and is based on several discordant published measurements. It does not have a very high weight. Exactly the same difficulties arise in obtaining the photographic and radio magnitudes, that is, an uncertainty as to the contribution from the faint outer regions.

TABLE 3

A COMPARISON OF THE RADIO AND PHOTOGRAPHIC MAGNITUDES OF SOME SOUTHERN NEBULAE

Nebula	Type	m_R	m_p	$m_R - m_p$
LMC	(M)	3.8	0.5	3.3
SMC	(M)	5.9	2.0	3.9
NGC 55	(M):	11.4	7.8	3.6
NGC 253	Sc	9.4	7.6	1.8
NGC 300	Sc	10.4	8.5	1.9
NGC 5236 (M83) ..	SBc	8.9	7.4	1.5
NGC 4945	SBc :	9.2	7.8	1.4
NGC 6744	SBbc	10.7	9.1	1.6
NGC 1068 (M77) ..	Sb	8.9	9.6	-0.7
I 5267	Sb	11.1	10.8	0.3
NGC 4594 (M104) ..	Sab	> 11.8	8.9	> 2.9
NGC 1291	SBo	> 11.6	9.5	> 2.1
NGC 3115	E7	> 11.9	10.15	> 1.7
47 Tucanae	Globular Cluster	> 11.0	3.0	> 8.0
NGC 362	Globular Cluster	> 10.8	6.0	> 4.8

The Small Magellanic Cloud

Again the photographic magnitude is an estimate of de Vaucouleurs (Buscombe, Gascoigne, and de Vaucouleurs 1954) and the same difficulties apply. However, in this case the uncertainty in the radio magnitude probably exceeds that in the photographic.

NGC 55

The astronomical data for this and the following have been taken from "A Revision of the Harvard Survey of Bright Galaxies" (de Vaucouleurs 1952-53), supplemented at times by additional information supplied by its author. In this Catalogue the photographic magnitudes are total magnitudes and are based on the best available standards. There is some difficulty in classifying NGC 55, as it is seen nearly edge-on, but the photographic evidence for classification as a Magellanic type appears quite strong and it has therefore

been adopted. As will be seen later this is consistent with the radio evidence. The source was near the limit of detectability and several observations were required to be sure of its existence. A sample record is shown in Figure 4 (a). The flux density is $1.8 \times 10^{-26} \text{ W m}^{-2} (\text{c/s})^{-1}$ and the radio magnitude has been calculated as before.

NGC 253

This is a relatively strong source and quite free of confusion so that reliable measurements are possible. The photographic magnitude is also reliable. A sample record is shown in Figure 4 (b).

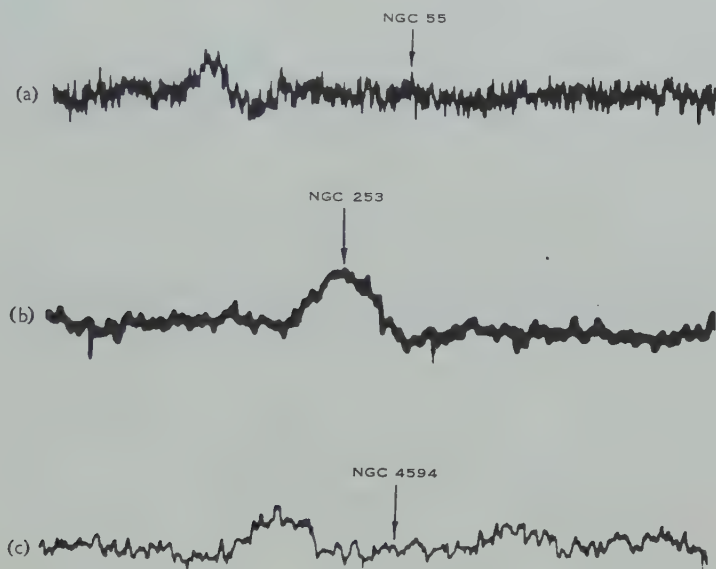


Fig. 4.—Records obtained on some bright galaxies. (a) A very weak radio source, NGC 55; (b) A relatively strong radio source, NGC 253; (c) The region around NGC 4594 (M104), a galaxy which could not be detected.

NGC 300

This galaxy is in a region of many weak radio sources and it is therefore difficult to estimate its radio emission or even to be sure of the identification. The radio source appears superimposed on a weaker source of large angular size and its apparent position is about 10 min of arc east of the galaxy, but the displacement could well be due to insufficient resolution. The photographic magnitude is also uncertain as the galaxy has a very low surface brightness and the published measures differ widely. The value quoted is an unpublished "best estimate" of de Vaucouleurs. Not very much weight can be given to the final value of $m_R - m_p$.

NGC 5236 (M83)

This galaxy gives a clear record without confusion but the photographic magnitude is of low weight.

NGC 4945

A good radio observation but the photographic magnitude does not carry much weight particularly as the galactic latitude is low and absorption effects uncertain. The galaxy is a late type seen nearly edge-on but the exact classification is in some doubt.

NGC 6744

The radio observations of the galaxy are good but the photographic magnitude has not much weight, particularly as the galactic latitude is relatively low.

NGC 1068 (M77)

Both the radio and optical observations of this galaxy are reliable. Although classified as an Sb it is of a fairly unusual type which shows strong and broad emission lines in the nucleus (Seyfert 1943). The relatively high value of radio emission is therefore not unexpected.

I 5267

An attempt was made to observe this comparatively faint galaxy primarily because a similarity in its appearance to that of NGC 5128 and NGC 1316 had been suggested by Evans (1949) and both the latter are now thought to be strong radio sources. The identification is not a good one as it is based on one observation only at an adjacent declination and there is some confusion from nearby sources. However, it is considered that there is a good chance that the identification is correct and the observation is in any case interesting as it demonstrates that the radiation is not markedly abnormal, if at all. Moreover, on plates taken by de Vaucouleurs there is no obvious similarity to NGC 5128.

NGC 4594 (M104)

This is a measurement of theoretical importance as M104 is a giant galaxy very rich in type II stars and also relatively bright. A sample record is shown in Figure 4 (c). Although there are several sources preceding and following the galaxy the region of its immediate vicinity shows no trace of excess emission. It is possible that a depression of the background radiation coincides with the position of the galaxy, but a more natural and likely interpretation is that the radio emission is too weak to be detected. Two particularly good records permit a very low upper limit to be set on its radio emission.

NGC 1291

This is one of the brighter early types accessible to the instrument, but the photographic magnitude is not known accurately and the value quoted is a "best estimate". There is no trace of radio emission.

NGC 3115

This is an almost completely dust free and probably gas free system and the photographic magnitude has a very high weight. Again there was no trace of radio emission on two consecutive high quality records.

Several northern galaxies have been detected by Brown and Hazard (1953). Their measurements are given in Table 4. One of their suggested identifications (NGC 5457) has been omitted as the position agreement was considered too poor, the discrepancy amounting to nearly 3° . The photographic magnitudes

TABLE 4

A COMPARISON OF THE RADIO AND PHOTOGRAPHIC MAGNITUDES OF SOME NORTHERN NEBULAE

Nebula				Type	m_R	m_p	$m_R - m_p$
NGC 5194-5 (M51)				Sc	9.7	8.5	1.2
NGC 224 (M31)				Sb	6.0	4.0	2.0
NGC 3031 (M81)				Sb	8.9	7.75	1.1
NGC 4258				Sb	9.8	9.1	0.7
NGC 2841				Sb	10.4	10.2	0.2
NGC 891				Sb	9.0	10.7	-1.7

are again derived from de Vaucouleurs' revision of the Shapley-Ames catalogue, and therefore differ from those quoted by Brown and Hazard who used the original catalogue.

The two sets of observations are combined in Figures 5 (a) and (b) in which, firstly, the radio magnitudes are plotted against the photographic, and secondly,

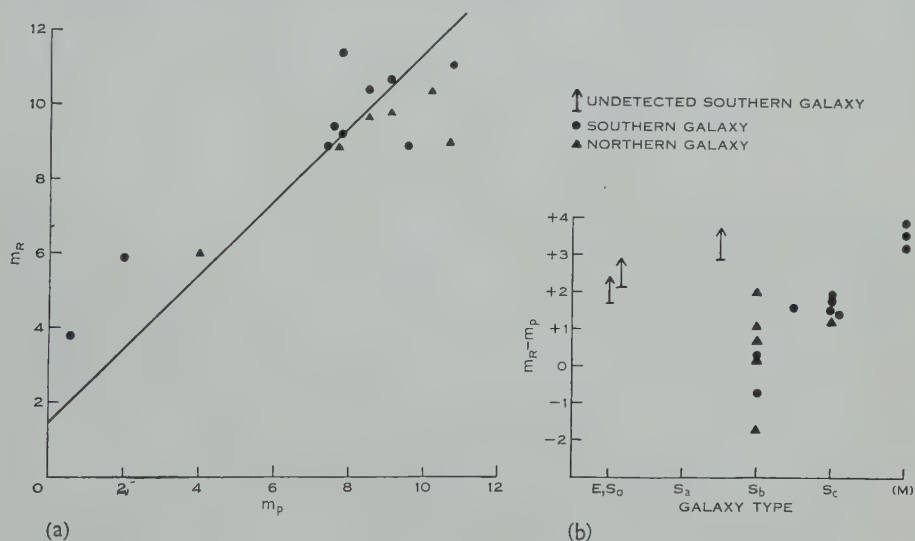


Fig. 5 (a).—A comparison of the radio and optical emission of all galaxies which have been detected.

Fig. 5 (b).—Illustrating the different radio emitting efficiencies of different galaxy types.

the differences between the two magnitudes are plotted against the type of galaxy. If all the galaxies had a constant ratio of radio to optical emission the points in Figure 5 (a) would lie on a straight line of slope 1.0. The line of this slope which fits the data best is shown; it corresponds to a mean value of

$m_R - m_p$ of $+1.4$, individual galaxies having a dispersion of 1.5 magnitudes. The northern galaxies fall mainly below this line and the southern galaxies mainly above it so that some calibration discrepancy might be suspected. However, there is no reason for supposing that galaxies of all types should have similar radio emitting efficiencies and it is desirable to compare the galaxy types separately.

In Figure 5 (*b*) where this is done it can be seen at once that there is no suggestion of any discrepancy. Unfortunately, the northern and southern galaxy types so far observed have little overlap, and significant cross checking of calibrations is not possible. However, it appears fairly safe to conclude that galaxies of intermediate type are more efficient emitters than both earlier and later types. The means of the various types, excluding the Sbc galaxy, together with their standard errors, are as follows:

$$\begin{aligned}\text{Sb,} \quad m_R - m_p &= +0.3 \pm 0.4, \\ \text{Sc,} \quad m_R - m_p &= +1.6 \pm 0.2, \\ (\text{M}), \quad m_R - m_p &= +3.6 \pm 0.2.\end{aligned}$$

The difference between Sb and Sc galaxies is hardly significant but the Magellanic types are very significantly weaker in radio emission than both. In view of earlier evidence concerning the type I distribution of emission in the Clouds, it is surprising that the Magellanic type galaxies which contain proportionally the greatest mass of type I population should be relatively weak radio emitters. This apparent anomaly will be discussed later after reviewing the evidence from the distribution of radio emission in our own galaxy. It is found to have a natural interpretation.

V. RADIATION FROM THE MILKY WAY

Although many attempts have been made to study the distribution of radio emission in the Galaxy, none has been completely successful. The difficulty is that at long wavelengths the aerial resolution has been insufficient to obtain unambiguous results while at short wavelengths the low receiver sensitivity and the low brightness of the non-thermal background component combine to restrict the useful information available. The resolution limitation imposed by a long wavelength, wide beam aerial has often been overlooked, but it is of vital importance because detail below a certain size is irretrievably lost (Bracewell and Roberts 1954). In one well-known survey (Bolton and Westfold 1950) this lower limit of size is nearly 8° . Some interferometer measurements of Scheuer and Ryle (1953) partly overcame this difficulty and gave effectively a high resolution at right angles to the galactic plane, but the poor resolution along the plane and the assumption of symmetry which had to be made produced results of doubtful quantitative value. Nevertheless, it is now found that these observations appear to give a good qualitative picture of the distribution at the longer wavelengths.

A complete survey of the southern sky is to be undertaken soon which, it is hoped, will give an accurate overall picture of the distribution, but, owing to the vast amount of detail which is discernible, it will be some time before the

results become available. Meanwhile it is possible, by obtaining a few galactic sections near the centre at fixed declinations, to check the various distributions already proposed. Three such sections are shown in Figure 6, and it can be seen at once that these do not fit the concept deriving from some of the earlier surveys, namely, that the spatial distribution is similar to that of the mass of the Galaxy (Westerhout and Oort 1951). Under such a hypothesis the width between half-brightness points should be about 17° . It appears instead that two galactic distributions are involved, one very sharply concentrated towards the galactic plane with a width between half-brightness points of about 3° , and less sharply but quite markedly towards the galactic centre; the other very broad

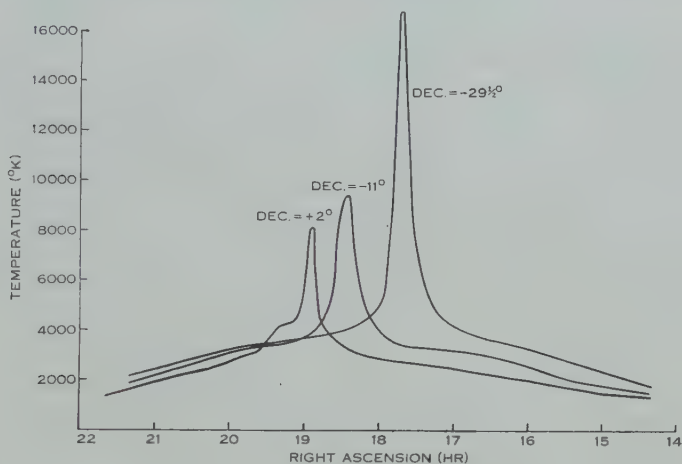


Fig. 6.—Some galactic sections near the centre. (Accuracy is about 10–20 per cent.)

and flat, with an estimated angle between half-brightness points (after allowing roughly for an isotropic extragalactic contribution) of 60 or 70° . The exact amount of extragalactic radiation cannot be estimated from these observations. It is no longer permissible to equate it to the temperature at the galactic poles which could be largely due to the very broad galactic distribution.

Such an overall picture has already been deduced by Shklovskii (1952, 1953) from an analysis of previous surveys. He maintains that the flattened distribution is due to thermal emission from ionized gas, and the broad distribution, which he concludes forms a spherical spatial distribution concentric with the galactic centre, to synchrotron type emission from relativistic electrons radiating in weak magnetic fields. Also from consideration of the contours of the Andromeda nebula obtained by Brown and Hazard (1951) he deduced a similar spherical distribution in that galaxy. The latter deduction has received support from some later measurements of Baldwin (1954).

The symmetry of the broad distribution of Figure 6 with respect to the galactic centre cannot be checked with a few observations because of the many and large brightness irregularities. However, two obvious properties of a spherical spatial distribution that might be roughly tested are that at a given

declination the sky brightness should, after excluding the contribution from the flattened subsystem, have its maximum value at the Right Ascension of the galactic centre and be a function only of the angular separation from the centre. Reference to Figure 6 shows that the sections are approximately consistent with these properties, and a few other observations of absolute temperatures which have been made also seem consistent.

It therefore appears safe to conclude that a division of the galactic radiation into the two subsystems proposed by Shklovskii can be assumed as a working hypothesis. It does not follow, however, that his interpretation of the mechanisms involved need be accepted; in fact it is quite easy to show that, contrary to his suggestion, the major contribution to the radiation from the flattened subsystem in the vicinity of the galactic centre must have a non-thermal origin. This follows immediately from a comparison with two higher frequency surveys, namely, those of Piddington and Minnett (1951) with a beam width of 2.8° at a wavelength of 25 cm, and by McGee and Bolton (1954) with a 2° beam at 75 cm wavelength. At a declination of $-29\frac{1}{2}^\circ$ these surveys yielded temperatures of 17°K at a wavelength of 25 cm,* and 150°K at a wavelength of 75 cm; at 3.5 m the temperature of the flattened subsystem at the same declination is $13,000^\circ\text{K}$. The apparent temperature spectrum of a mass of thermally emitting gas is given by $T \propto \lambda^n$, where n lies between 0 and 2, depending on the opacity. The temperature at 3.5 m is far too high to be fitted by this law which would predict a temperature less than 3400°K at that wavelength. However, comparisons between the two other measurements suggest that the radiation at shorter wavelengths is predominantly due to thermal emission from the ionized gas. By means of similar comparisons it is easy to demonstrate that the spherical subsystem also has a non-thermal spectrum.

The thickness of the flattened subsystem near the galactic centre is about 400 parsecs, corresponding to the observed angular width of 3° . This is comparable with the thickness of the neutral hydrogen layer, which suggests a relation to the Population I component of the Galaxy. A similar estimate for the spherical subsystem is necessarily very crude. The value obtained, which depends greatly on the form of the radial distribution, is of the order of 10 kiloparsecs. The radio emission undoubtedly extends much further in an attenuated form. It is not possible to decide at present whether the distribution is oblate, or truly spherical. The observations tend to favour the first possibility.

VI. DISCUSSION

The radio emitting properties of a selection of galaxies have now been examined. Data have been presented showing the total emission of considerable

* The discrete source near the galactic centre has been ignored in estimating the observed temperature of Piddington and Minnett because it is quite clear from several other observers' results that its declination is approximately $-28\frac{1}{2}^\circ$, i.e. well removed from the comparison declination of $-29\frac{1}{2}^\circ$. The declination scale of Piddington and Minnett appears to be in error by about $1\frac{1}{2}^\circ$.

numbers of relatively bright galaxies, the total emission and the distribution of radiation in the Magellanic Clouds and some features of the distribution of radio emission in the Milky Way. It is interesting to speculate on the interconnexion between the sets of observations and to attempt to explain the data by a general theory of the radio emission of a "normal" galaxy.

It has been shown that the galactic radiation can be assumed to originate in two distinct subsystems, one highly flattened and apparently related to the type I population of the Galaxy and the other, roughly spherical, concentric with the centre of the Galaxy and very diffuse, forming a corona surrounding the system. One or the other of these subsystems is recognizable in other galaxies also. In the Clouds the Population I type appears to predominate, as the distribution of radio emission is similar to that of the gas and early type stars. However, the almost plan view of the systems precludes a direct measurement of the thickness of the distribution. In M31 and possibly on a very much smaller scale in the Small Cloud the spherical subsystem can be recognized. In both the Galaxy and M31 this spherical system must be the major source of radiation since in the former it is easily seen that its integrated emission far exceeds that of the sharply concentrated type I component even though the data are insufficient to compute the ratio with any accuracy;* similarly the contours of M31 obtained by Brown and Hazard show no evidence of a flattened subsystem so that, if present, it must be weak. A natural assumption is that the same two subsystems are present in a greater or lesser degree in all normal galaxies. The relatively high emission of intermediate type galaxies may then be interpreted as due to the possession of a more developed "corona" than either earlier or later types.

It is interesting to observe that the "abnormal" radio galaxy NGC 5128 also displays two subsystems of radio emission, one associated with the band of dust crossing the nucleus and the other forming an extensive corona surrounding the system (Mills 1953).

The flattened subsystem has two basic components. One is due to free-free emission from ionized hydrogen: as the brightness due to this emission process is independent of the wavelength when the opacity is small, it becomes important at the shorter wavelengths. This emission process in the Galaxy has been discussed by many authors, e.g. Reber (1940), Piddington (1951), Westerhout and Oort (1951), Shklovskii (1952, 1953), Haddock, Mayer, and Sloanaker (1954), and many others. The other component is of non-thermal origin and increases in brightness at the longer wavelengths to dominate the picture entirely. Identifications which have been made between some strong radio sources and peculiar galactic nebulae (see for instance a summary by Pawsey (1955)) suggest the possibility that the integrated emission from similar nebulae may be the main source of the non-thermal component in the Galaxy. If so, it is possible that it cannot be classed as purely a Population I type of distribution because

* A rough estimate of the ratio is about 5 or 10.

included among these identifications are the nebulous remains of supernovae, and it is thought that these may be associated with Population II.*

The spherical subsystem, while superficially resembling a pure Population II system, lacks the high central concentration which they display, and in fact, as pointed out by Shklovskii, is quite unlike any known optical feature of the Galaxy. In addition radio emission is absent from the type II stars of globular clusters and a relatively low upper limit can be set to the emission from early type galaxies which have a high proportion of Population II. It therefore appears that this subsystem is unrelated to any distribution of Population II stars, and in general the total radiation of a galaxy at the longer wavelengths cannot be ascribed to a subclass of objects belonging to Baade's Populations I or II.

Turning now to speculation concerning the mechanism of non-thermal radio emission, it has been suggested by Hoyle (1954) and Twiss (1954) that some of the galactic radio nebulae which are associated with highly turbulent ionized gas may represent localized regions containing high energy electrons accelerated by the turbulent gas in a Fermi type process to relativistic energies (Fermi 1949). The relativistic electrons would then radiate at radio frequencies in the concomitant magnetic fields. The observed properties of such nebulae are readily explicable in this way and the necessary physical conditions are reasonable. It is natural to attempt to extend this concept to the spherical subsystem also.

It has been shown by Pikelner (1953) that magnetic fields are likely to exist in the highly rarefied medium between gas clouds with approximately the same average strength as they exhibit inside the clouds. From equipartition considerations this implies that the random macroscopic velocities of the rarefied medium are higher than those of the gas clouds by a factor of the order of 10. It is easily shown that, because of its high velocity, this rarefied medium and its associated magnetic fields must tend towards a spherical spatial distribution. Shklovskii (1953) has suggested that the source of the spherical subsystem of radio emission is the radiation from relativistic electrons moving in this spherical field system. However, he ignores the dynamic effect of the turbulent medium in accelerating charged particles and consequently considers the relativistic electrons to be introduced at their full energy from an external source, in fact, from novae and supernovae. While it is possible that these could be a source of high energy electrons, it appears probable that the electron energy would be increased or at least maintained against radiation losses by the energy of the

* Shklovskii (1953) maintains that all non-thermally emitting nebulosities are the remains of supernovae, but Baade and Minkowski (1954), basing their argument on the frequencies of supernovae observed in external galaxies, conclude that such objects must be comparatively rare and, if their figures are accepted, could account for only an insignificant proportion of the total observed brightness. It is hoped that a current series of radio observations of some southern emission nebulae will help clarify the matter.

turbulent gas which would most likely be the dominant factor in determining the total radio emission.*

This theory has attractive features and leads to a simple explanation of the observed differences in the radio emission of different classes of galaxies. Early type galaxies, comparatively lacking in interstellar gas, are deficient in the corona of gas and magnetic fields; consequently the emission is low. Late type galaxies of the Magellanic type, on the other hand, possess abundant supplies of gas but do not have sufficient gravitational energy to retain a high velocity corona. Intermediate types possess a good supply of gas from their type I population and a high gravitational energy because of their type II population and therefore radiate strongly.†

The operation of this radiation process in the Galaxy may be investigated quantitatively to determine the physical conditions required. First we may estimate the mean energy of relativistic electrons by equating the rate of gain of energy due to "magnetic collisions" (Fermi 1949) to the rate of loss by radiation (Schwinger 1949), thus,

$$\left(\frac{V_c}{c}\right)^2 \cdot E \cdot \nu_c \sim 1.6 \times 10^{-15} H^2 \frac{E^2}{(m_0 c^2)^2}, \quad \dots\dots\dots (1)$$

where V_c is the turbulent velocity of the corona, ν_c is the "magnetic collision frequency" in the corona, E is the electron energy, and H the magnetic field. The latter may be obtained from the following relation:

$$\frac{1}{8\pi} \bar{H}^2 = \frac{1}{2} \rho_c \bar{V}_c^2, \quad \dots\dots\dots (2)$$

where ρ_c is the density of the coronal gas.

Combining (1) and (2) and inserting some physical constants we have

$$E \sim 3 \times 10^{-7} \frac{\nu_c}{\rho_c} \text{ eV}. \quad \dots\dots\dots (3)$$

* In his theory of particle acceleration Fermi (1949) found that the rate of gain of energy of electrons is insufficient to counterbalance the loss by ionizing collisions with hydrogen atoms. However, he was considering regions very close to the galactic plane where dense neutral hydrogen clouds abound and random velocities are low. If a corona of tenuous high velocity gas does exist, the rate of gain of energy, which is proportional to the square of the velocity, must be increased by several orders of magnitude. It also seems possible that such a gaseous envelope would, at least in the Galaxy, be largely ionized because the shielding effect of the dense hydrogen clouds would be small in directions normal to the plane. The radius of a Stromgren sphere in a medium of density 10^{-2} atoms cm^{-3} surrounding an early type star is of the order of several thousand parsecs (Stromgren 1939); so that ionization is possible at large distances from the plane. Collisional ionization is also a possibility. Such an ionized corona would be unobservable optically since its emission measure would be only of the order of 10 to 100.

† An attempt has been made to extend this idea a little further by comparing the radio emission of a galaxy with the kinetic energy of its corona. In a few cases the kinetic energy has been estimated by assuming a mass proportional to the Population I component and a velocity determined by the total gravitational energy. By adopting reasonable values for these parameters a strict proportionality between kinetic energy and radio emission is obtainable. However, uncertainties in the astronomical data are so large at present that no positive conclusions can be drawn.

Plausible values of v_c and ρ_c are $10^{-10} \text{ sec}^{-1}$ and $3 \times 10^{-26} \text{ g cm}^{-3}$ respectively, yielding an average electron energy of 10^8 eV .

By using some formulae given by Ginsberg (1951) the frequency of maximum radio emission ($\nu_{\text{max.}}$), may be calculated, and also the electron density n required to produce the observed radio emission at that frequency, thus

$$\nu_{\text{max.}} = 5 \times 10^{-5} H E^2, \dots\dots\dots (4)$$

$$n = 10^{23} \frac{kT}{HD} \left(\frac{\nu_{\text{max.}}}{c} \right)^2, \dots\dots\dots (5)$$

where D is the distance through the corona in the direction the measurement is made and T is the brightness temperature of the radio emission.

The magnetic field H may be estimated from equation (2) using the value adopted earlier for ρ_c and a velocity of 200 km/sec . The velocity has been obtained using the virial theorem which seems permissible since the corona is approximately spherical. The resulting field is of the order of 10^{-5} gauss which, when substituted in equation (4), yields a frequency of maximum emission of the order of 5 Mc/s .

No observations are available at this frequency, the nearest being those of Higgins and Shain (1954) at a frequency of 9.5 Mc/s . If their results are extrapolated to the lower frequency a brightness temperature of about $5 \times 10^6 \text{ }^\circ\text{K}$ is obtained near the galactic centre. Taking D equal to $6 \times 10^{22} \text{ cm}$ we then find from equation (5) that the electron density is of the order of 10^{-11} to 10^{-12} electrons cm^{-3} .

All the calculated values seem very plausible and do not appear to contradict any observational results. We may therefore conclude that the mechanism is a possible one.

VII. CONCLUSIONS

It appears that the major portion of the radiation of "normal" galaxies at metre wavelengths can be explained qualitatively and perhaps even quantitatively by emission from relativistic electrons radiating in weak interstellar magnetic fields. It should be emphasized, however, that this conclusion does not follow as a *necessary* consequence of the data. For this reason the theory has been developed only in a very sketchy form.

On somewhat firmer ground is the conclusion that radio emission can occur with two separate and distinct distributions, one displaying characteristics of Baade's Population I, and the other having a diffuse roughly spherical spatial distribution forming a "corona" around a galaxy not, however, related to the corona of globular clusters often observed. Both distributions have a non-thermal spectrum and the relative contributions of each vary between different galaxy types. Very late type galaxies such as the Magellanic Clouds possess a high proportion of the Population I type while early type galaxies are deficient in both. Intermediate type galaxies, by virtue of an extensive "corona", are the most efficient radio emitters.

VIII. ACKNOWLEDGMENTS

These results were made possible by the work of Mr. A. G. Little and Mr. K. V. Sheridan in constructing and maintaining the aerial and its associated equipment and in making many of the observations described. The author is also indebted to Dr. G. de Vaucouleurs for much advice about the astronomical data and for the generous communication of many of his own observational results prior to publication.

IX. REFERENCES

- BAADE, W., and MINKOWSKI, R. (1954).—*Astrophys. J.* **119**: 206.
- BALDWIN, J. E. (1954).—*Nature* **174**: 320.
- BOLTON, J. G., and WESTFOLD, K. C. (1950).—*Aust. J. Sci. Res. A* **3**: 19.
- BRACEWELL, R. N., and ROBERTS, J. A. (1954).—*Aust. J. Phys.* **7**: 615.
- BROWN, R. H., and HAZARD, C. (1951).—*Mon. Not. R. Astr. Soc.* **110**: 357.
- BROWN, R. H., and HAZARD, C. (1952).—*Phil. Mag.* **43**: 137.
- BROWN, R. H., and HAZARD, C. (1953).—*Mon. Not. R. Astr. Soc.* **113**: 123.
- BUSCOMBE, W., GASCOIGNE, S. C. B., and DE VAUCOULEURS, G. (1954).—*Aust. J. Sci.* **17** Suppl.: 1.
- EVANS, D. S. (1949).—*Mon. Not. R. Astr. Soc.* **109**: 94.
- FERMI, E. (1949).—*Phys. Rev.* **75**: 1169.
- GINSBURG, V. L. (1951).—*Dokl. Akad. Nauk, SSSR* **76**: 377.
- HADDOCK, F. T., MAYER, C. H., and SLOANAKER, M. (1954).—*Nature* **174**: 176.
- HENIZE, K. G., and MILLER, F. D. (1951).—*Publ. Obs. Univ. Mich.* **10**: 75.
- HIGGINS, C. S., and SHAIN, C. A. (1954).—*Aust. J. Phys.* **7**: 460.
- HOYLE, F. (1954).—*Nature* **173**: 483.
- KERR, F. J., HINDMAN, J. V., and ROBINSON, B. J. (1954).—*Aust. J. Phys.* **7**: 297.
- KERR, F. J., and DE VAUCOULEURS, G. (1955).—*Aust. J. Phys.* **8** (in press).
- MCGEE, R. X., and BOLTON, J. G. (1954).—*Nature* **173**: 985.
- MILLS, B. Y. (1953).—*Aust. J. Phys.* **6**: 452.
- MILLS, B. Y., and LITTLE, A. G. (1953).—*Aust. J. Phys.* **6**: 272.
- PAWSEY, J. L. (1955).—A catalogue of reliably known discrete sources of cosmic radio waves. *Astrophys. J.* **121**: 1.
- PIDDINGTON, J. H. (1951).—*Mon. Not. R. Astr. Soc.* **111**: 45.
- PIDDINGTON, J. H., and MINNETT, H. C. (1951).—*Aust. J. Sci. Res.* **4**: 459.
- PIKELNER, S. B. (1953).—*Dokl. Akad. Nauk, SSSR* **88**: 229.
- REBER, G. (1940).—*Proc. Inst. Radio Engrs. N.Y.* **28**: 68.
- SCHEUER, P. A. G., and RYLE, M. (1953).—*Mon. Not. R. Astr. Soc.* **113**: 3.
- SCHWINGER, J. (1949).—*Phys. Rev.* **75**: 1912.
- SEYFERT, C. K. (1943).—*Astrophys. J.* **97**: 28.
- SHKLOVSKII, I. S. (1952).—*Astr. J., Moscow* **29**: 418.
- SHKLOVSKII, I. S. (1953).—*Astr. J., Moscow* **30**: 15.
- STROMGREN, B. (1939).—*Astrophys. J.* **89**: 526.
- TWISS, R. Q. (1954).—*Phil. Mag.* **45**: 249.
- U.R.S.I. (1954).—Discrete sources of extraterrestrial radio noise. Spec. Rep. No. 3.
- DE VAUCOULEURS, G. (1952–53).—“A Revision of the Harvard Survey of Bright Galaxies.” Australian National University Monograph.
- DE VAUCOULEURS, G. (1954).—*Observatory* **74**: 158.
- DE VAUCOULEURS, G. (1955a).—Studies of the Magellanic Clouds. I. Dimensions and structure of the Large Cloud. *Astr. J.* (in press).
- DE VAUCOULEURS, G. (1955b).—Studies of the Magellanic Clouds. II. Dimensions and structure of the Small Cloud. *Astr. J.* (in press).
- WESTERHOUT, G., and OORT, J. H. (1951).—*B.A.N.* **10**: 323.

THE RADAR DETERMINATION OF METEOR SHOWERS IN THE SOUTHERN HEMISPHERE

By C. D. ELLYETT* and K. W. ROTH*

[*Manuscript received May 20, 1955*]

Summary

During the greater part of 1953 a radar survey was made of meteor activity in the southern hemisphere. The results are presented, together with a brief description of the apparatus used. Radiants are calculated from daily range-time plots of meteor echoes, and the resulting showers are in accord with data reduced from earlier visual observations.

Both radar and visual data for the southern hemisphere show many night radiants between mid June and mid August. Showers in this period are clearly defined but overlap in date, and there is a marked drift in the direction of the meteor activity as the date progresses. Two new southern hemisphere daylight showers of moderate strength have been found by radar in June and October respectively, and considerable confused night activity is present in early December. The greater proportion of both the northern and the southern hemisphere meteor shower radiants appear to lie close to the plane of the ecliptic.

I. INTRODUCTION

Meteor showers have been studied extensively in the northern hemisphere (Almond, Bullough, and Hawkins 1952 ; Hawkins and Almond 1952*a*, 1952*b* ; Bullough 1954 ; Prentice 1955). Most showers contain many individual meteors of sufficient size to leave visual trails in the upper atmosphere, and recur annually at fixed dates. The meteors in any given shower travel in a definite orbit about the Sun, and this orbit generally intersects the Earth's orbit over a number of days, determined by the width of the stream at the chord of the Earth's crossing. Knowledge of individual meteor shower directions in space, or radiants, seen from the northern hemisphere is fairly complete, but the southern hemisphere radiants, which have been based on comparatively few visual observations, are much less well known.

The aim of the present paper is to indicate the technique and present the results of a radar survey of meteor activity in the southern hemisphere carried out in 1953. These results are compared with earlier visual observations which are outlined in Section II. Finally, when the principal northern and southern hemisphere meteor radiants are plotted together it is shown that many of the major meteor showers lie closely in the plane of the ecliptic.

II. MOST PROBABLE SOUTHERN HEMISPHERE METEOR SHOWERS

Lists of southern hemisphere visual meteor observations have been published by Hoffmeister (1948) and by McIntosh (1934, 1935, 1936, 1940). Radiants found from these visual observations were based on very small numbers of

* Canterbury University College, Christchurch, N.Z.

meteors, and in consequence many sporadic groupings have undoubtedly been regarded as radiants. Sporadic meteors from random directions are always present, and no criterion has been established which specifies the number or the rate of arrival of meteors from a particular direction which can be regarded as constituting a stream. A radiant due to McIntosh requires a minimum of four similarly directed meteor paths on a single night. Separate radiants are recorded on separate nights and on separate years. These results have been reduced to Table 1, in which showers seen by McIntosh are only included if nine or more separate but coincident radiants have been listed. Hoffmeister gives a probability rating from 1 to 10 for observed radiants, and of these only probability 10 has been included in Table 1. Visual showers recorded on a single night only have been excluded independent of rating.

A complete radar study of the meteor streams incident in the northern hemisphere has been made by the Jodrell Bank research team of the University of Manchester (Almond, Bullough, and Hawkins 1952; Hawkins and Almond 1952*a*, 1952*b*; Bullough 1954), and their results have been shown to be in accord with the less complete early visual work. Northern hemisphere radar experience indicates that showers of lower altitude than about 20° above the horizon at transit are only weakly recorded. In considering northern hemisphere radar results for possible inclusion in Table 1 all showers for which the mean radiant exceeds 26.5° N. have therefore been excluded, as the altitude of such showers at transit is less than 20° above the horizon at Christchurch, New Zealand (43.5° S.). Mid northern and mid southern hemisphere observations overlap in an equatorial belt.

The primary purpose of the present paper has been to identify showers. The accuracy of the Christchurch radar observations does not yet justify a determination of the slight day-to-day changes in radiant as the Earth moves through a particular shower. Consequently where such variations have been reported elsewhere, they have been grouped as mean radiants in Table 1.

III. EQUIPMENT AND METHOD

Surplus defence radar equipment provided the basic transmitting and receiving units. Pulses of 75 kW peak power and $3.5 \mu\text{sec}$ width were generated by the transmitter at a frequency of 69 Mc/s and a recurrence rate of 145/sec. Automatic control was provided, so that reswitching was effected after transient overloads or failures in the mains power supply. The equipment was kept inoperative between 14 and 21 hr L.T. each day in order to conserve power and film when the rate was at its minimum.

The aerial array consisted of 12 horizontal half-wavelength dipoles, all lying in the one vertical plane and end-fed in pairs, with each vertical stack containing three elements. Fine gauge wire netting, located one-eighth wavelength behind the plane of the dipoles, served as a reflecting screen, and the centre of the array was located one wavelength above the ground. The measured and calculated horizontal radiation patterns agreed very closely. The vertical pattern, which could not readily be measured, was calculated on the assumption of perfect ground reflection. The resulting cone of radiation has maximum

power at 15° elevation, with beam widths between half-power points of 14° and 20° respectively in the vertical and horizontal directions. An aerial power gain of 75 is obtained compared with a half-wave dipole. Small side lobes are situated on either side of, and at the same elevation as the main lobe, but the power gain is only 4.4 per cent. of the maximum. Smaller lobes present at greater elevations

TABLE 1

OCCURRENCE OF SOUTHERN HEMISPHERE METEOR SHOWERS ESTIMATED FROM SOUTHERN HEMISPHERE VISUAL DATA AND NORTHERN HEMISPHERE RADAR DATA

Shower Number	Date*	Apparent Radiant		Constellation	Observer†	Number of Radiants, or Rating
		R.A.,	Dec.			
	Jan. Feb. Mar. }	Nil				
I	Apr. 1-21	228/239°,	-19/-22°	♎ Libra	M (V)	18
II	Apr. 4-30	208½°,	-10°	♍ Virgo	M (V)	24
III	Apr. 28-May 16	330½/346½°,	+3½/-3°	♊ Aquarius	M (V)	24
	May 1-6	313/338°,	+3/-4°		J (R)	A
IV	Apr. 29-May 7	323½°,	-17°	♐ Capricornus	M (V)	9
V	May 29-June 19	38/50°,	+16/+30°	♈ Aries	J (R)	A
VI	June 1-16	57/72°,	+14/+30°	♉ Perseus	J (R)	A
VII	June 1-10	250/259°,	-22/-23°	♏ Ophiuchus	M (V)	9
VIII	June 25-28	196/199°,	-17/-23°	♐ Hydra	H (V)	Py. 10
IX	June 26-July 4	81/92°,	+13/+26°	♉ Taurus	J (R)	A
X	June 30-July 10	262/275°,	-20°	♐ Sagittarius	M (V)	10
XI	July 2-19	283°,	-15°	Sagittarius	M (V)	9
XII	July 5-9	257/261°,	-8/-12°	♐ Ophiuchus	M (V)	9
XIII	July 5-8	238°,	-19°	♏ Scorpius	M (V)	9
XIV	July 10-Aug. 5	300/325°,	-10/-19°	♐ Capricornus	M (V)	63
XV	July 22-Aug. 15	331/351°,	0/-17°	♊ Aquarius	M (V)	78
	July 29-31	345°,	-20°		H (V)	Py. 10
	July 25-Aug. 6	339°,	-12°		L (R)	
		340±5°,	-17±2°		K (R)	
XVI	July 19-Aug. 1	352/11°,	-18°	♐ Cetus	M (V)	13
XVII	July 14-Aug. 22	330/339°,	-30°	♐ Pisces	M (V)	11
				Australis		
XVIII	July 26-Aug. 8	337/350°,	-30/-33°	♐ Pisces	M (V)	24
				Australis		
XIX	Oct. 15-25	97°,	+11°	♈ Orion	J (R)	
XX	Oct. 26-Nov. 16	58°,	+23°	Taurus	J (R)	
XXI	Nov. 17-18	150/151°,	+21/+22°	♈ Leo	H (V)	Py. 10
	Nov. 15-20	152°,	+22°		B (R)	
	Dec.	Nil				

* Dates in this table are based on N.Z.S.T.=U.T.+12 hr.

† (V)=Southern hemisphere visual observations, (R)=northern hemisphere radar observations. M=McIntosh (1934, 1935, 1936, 1940); J=Jodrell Bank group, University of Manchester (Almond, Bullough, and Hawkins 1952; Hawkins and Almond 1952a, 1952b; Bullough 1954); H=Hoffmeister (1948); L=Lindblad (1952); K=McKinley (1954); B=British Astr. Ass. Handbook, 1955.

can only give rise to short range echoes. The effect of the side lobes is thus slight, and has not been further considered. In order to determine radiants the array, which was rotatable in azimuth, was held at 22.5° N. of W. and 22.5° S. of W. respectively on alternate days.

Before observations were commenced a new input stage was built for the receiver which had been supplied, giving an overall receiver noise figure of 6.8 db. The bandwidth was 300 kc/s. The system sensitivity was adjusted so that the non-shower rate lay roughly between 5 and 10/hr throughout the year. The figure of 8/hr given by McIntosh (personal communication) for visual viewing in New Zealand with a 50° viewing cone is therefore somewhat less than the radar rate, when the smaller radar viewing angle is taken into account.

Intensity modulation was used to present the receiver output signal on the oscillograph, which was viewed by 35 mm film moving at 1.23×10^{-2} cm/sec. A continuous record was thus obtained showing echo range transversely and time longitudinally. (A small step voltage was added on alternate sweeps across the screen, causing genuine echoes to appear as close dense twin spots or lines, against the random spots resulting from background noise.)

A radiant passing through the aerial beam is characterized by a rise in echo rate and a progressive increase in echo range to a maximum, followed by a fairly rapid collapse (Clegg 1948). Some 50,000 echoes, recorded during the survey, were plotted on daily range-time graphs. Transparencies of all possible radiant envelopes were drawn independently by both authors, following the Clegg technique (Clegg 1948; Aspinall, Clegg, and Hawkins 1951). The reality of the various envelopes was considered, and similarly shaped envelopes were then looked for on adjacent days, since two days with different aerial azimuth setting are necessary before a radiant can be calculated. The remaining radiant results calculated from these envelopes, after all dubious envelopes or results had been discarded, were then grouped where possible into adjacent pairs of days with similar radiants. Interlocking radiants and subgroups made the whole procedure formidable but not impossible. (Envelopes of meteors referring to possible showers from the direction of the South pole could not be delineated.)

IV. SOUTHERN HEMISPHERE RADAR SHOWER RESULTS

The reproducibility of results is indicated in Figure 1. The top two plots are for adjacent nights with the same aerial setting, and the third plot is for the same night and aerial setting as the second plot, but in the following year when the sensitivity had been adjusted slightly below that of the 1953 records. Echoes in general are only observed when the aerial beam is at right angles to the incoming meteor trails (Clegg 1948). In all three cases a major group passes through the aerial beam with maximum range at approximately 15 hr. Another group is indicated at 13 hr. Mid-morning activity is also pronounced, with a group present at 21 hr and another probably present at 22 hr 20 min. During the first half of June, for other reasons, the aerial was not turned, so radiants could not be calculated during this period, but the activity—both day and night—outweighed that of any other month.

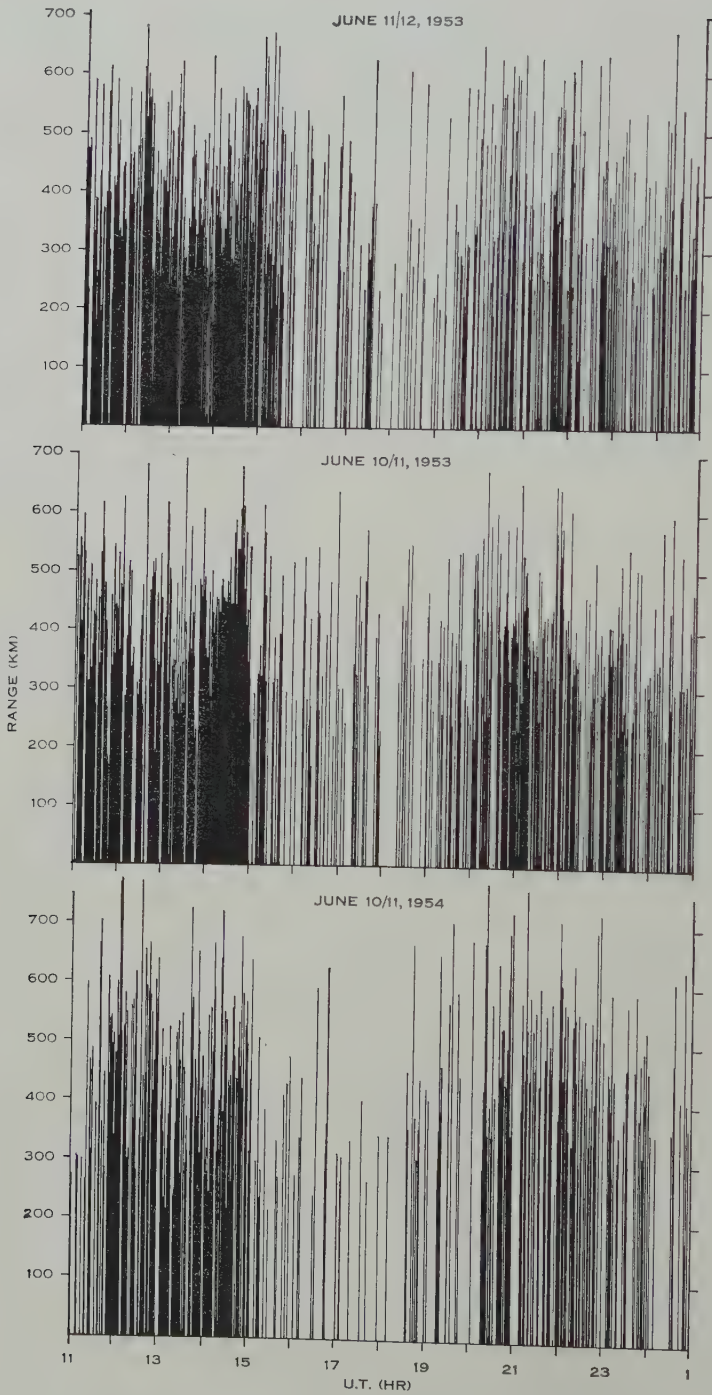


Fig. 1.—Meteor echoes, received with directional aerial array, showing strong day and night activity.

For comparison a September plot with unchanged sensitivity is shown in Figure 2. During this month echoes were scattered, and no radiant could be identified which lasted for more than two adjacent nights.

Each radiant group, found as described in Section III, was given an arbitrary grading. This was based on the consistency of the shape, range, and strength of the group, combined with factors incorporating the number of envelopes seen, the number of pairs calculated, and the compactness of the radiant limits. Four grades were made. Grade A is considered to be certain, and grade B is highly probable. The final results obtained are set out in Table 2.

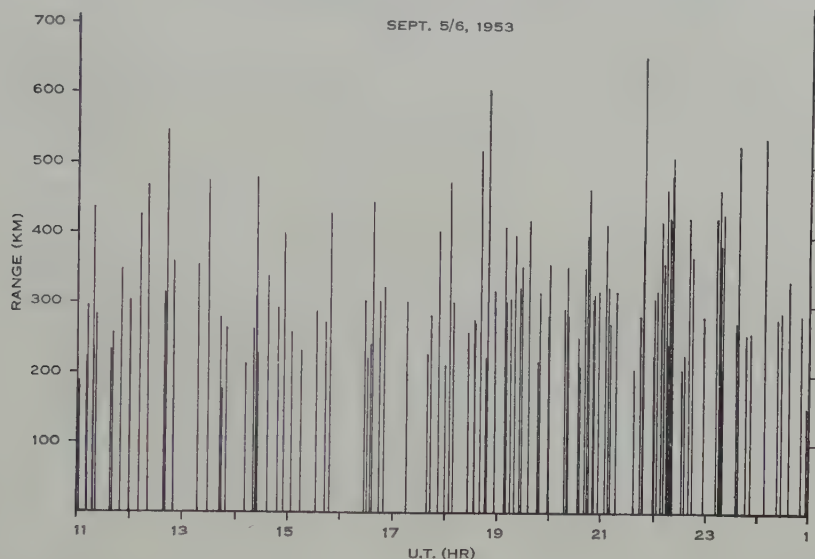


Fig. 2.—Meteor echoes received during an inactive period. (Mean rate=8.6/hr.)

A comparison between the most probable southern hemisphere meteor showers evaluated from visual data (Table 1) and from the radar results (Table 2) is given in Figure 3. Showers shown in contact are considered to be subgroups of a single major radiant. In general the radar activity associated with any given shower covers fewer nights than the visual estimate, although this is only to be expected since the radar technique is more stringent in its selection. Showers of similar radiant determined by the two methods are connected by cross lines. The overlap of dates is often incomplete. However, there is a sufficiently high degree of correspondence to identify the radar observations clearly with the earlier visual results.

Of the 21 visual showers listed in Table 1, only numbers IV, V, VI, VII, VIII, XIII, XIX, XX, and XXI are not immediately evident by radar. Of these, numbers V, VI, XIX, XX, and XXI are essentially northern hemisphere radiants. Some radar trace of the Leonids (XXI) was found, but was considered insufficient to include in Table 2. The Leonids (XXI), Taurids (XX), and Orionids (XIX) are known to be weak, and their absence at an observatory situated at 43.5°S . is not surprising. Numbers V and VI are very strong

TABLE 2

SOUTHERN HEMISPHERE METEOR SHOWERS CALCULATED FROM RADAR METEOR ECHOES

Shower Number	Date*	No. of Envelopes Included	Mean Apparent Radiant		Limits of Radiant		Constellation	Arbitrary Grading of Shower Probability
			R.A.	Dec.	R.A.	Dec.		
1	Feb. 17-22	6	297°, -6°		296/298°, -1/-10°		Aquila	D
2	Apr. (10)-(15)	5	212°, -9°		208/214°, 0/-14°		Virgo	C
3	Apr. 29-May 8	9	228°, -23°		226/229°, -19/-24°		Libra	C
4	Apr. 29-May 4	5	343°, +1°		342/344°, +2/-2°		Pisces/ Aquarius	D
5	June (17)-29	12	303°, -21°		302/305°, -19/-23°		σ Capri- cornus	B
6	June 19-25	6	268°, -23°		268/269°, -23°		μ Sagittarius	C
7	June (17)-28	11	270°, -13°		267/274°, -10/-15°		Ophiuchus/ Aquila	A
8	June (17)-28	10	263°, -21°		260/266°, -19/-23°		ξ Ophiuchus	A
9	June (17)-21	5	252°, -20°		249/254°, -19/-21°		η Ophiuchus	D
10	June (17)-26	9	84°, +23°		82/85°, +20/+25°		β/ζ Taurus	B
11	June 19-27	8	69°, -27°		65/71°, -23/-30°		Eridanus	B
12	June 18-21	4	48°, -25°		46/50°, -21/-26°		Eridanus	D
13	July 3-13	9	302°, -33°		301/303°, -32/-34°		θ Sagittarius	B
14	June 29-July 13	13	300°, -19°		299/301°, -18/-20°		Sagittarius	B
15	July 5-15	11	310°, -33°		309/311°, -26/-38°		Capri- cornus/ Micro- scopium	A
16	July 3-9	5	317°, -17°		316/318°, -15/-18°		θ Capri- cornus	D
17	July 17-25	8	275°, -8°		271/279°, -5/-11°		α Aquila	C
18	July 16-25	10	321°, +2°		319/322°, +1/+2°		β Aquarius	A
19	July 20-Aug. 2	12	299°, -3°		297/302°, +1/-7°		θ Aquila	A
20	July 21-30	8	322°, -17°		320/322°, -14/-19°		δ Capri- cornus	C
21	July 21-26	5	328°, -27°		326/329°, -26/-27°		Pisces Australis	B
22	July 30-Aug. 5	7	338°, -18°		334/338°, -14/-22°		δ Aquarius	A
23	July 29-Aug. 12	14	301°, -10°		297/304°, -7/-14°		α Capri- cornus	A
24	Aug. 1-12	13	321°, -8°		319/323°, -2/-11°		ξ Aquarius	B
25	Aug. 2-9	8	329°, -11°		328/331°, -9/-14°		θ Aquarius	C
26	Oct. 7-14	6	63°, -8°		62/65°, -8/-9°		ξ Eridanus	D
27	Oct. 9-16	8	49°, +3°		48/50°, 0/+6°		Cetus	D
28	Oct. 23-31	8	46°, 0°		44/48°, +1/-3°		α Cetus	C
29	Oct. 22-28	7	200°, -19°		199/201°, -10/-24°		Virgo/ Hydra	B
30	Oct. 31-Nov. 7	7	204°, -6°		202/205°, -1/-8°		α Virgo	C

* Dates in this table are based on N.Z.S.T. Dates shown in brackets represent activity present immediately before or after a non-operative period. Such showers may have started or terminated earlier or later respectively than the date shown.

TABLE 2 (Continued)

Shower Number	Date*	No. of Envelopes Included	Mean Apparent Radiant		Limits of Radiant		Constellation	Arbitrary Grading of Shower Probability
			R.A.	Dec.	R.A.	Dec.		
31	Nov. 12-(16)	5	78°, -15°		76/79°, -11/-19°		μ Lepus	D
32	Dec. (1)-8	5	224°, -27°		222/228°, -25/-29°		σ Libra	C
33	Dec. (1)-10	7	147°, -5°		147°, -1/-9°		γ Sextans	D
34	Dec. 2-11	6	253°, -20°		252/256°, -18/-23°		Ophiuchus	D
35	Dec. (1)-11	6	235°, -19°		234/235°, -15/-22°		θ Libra	D
36	Dec. 2-10	6	114°, -25°		112/116°, -24/-26°		Puppis	C
37	Dec. (1)-6	5	88°, -3°		86/91°, -2/-4°		Orion	D
38	Dec. 12-19	5	112°, -16°		109/114°, -14/-18°		Puppis	D
39	Dec. 12-20	8	193°, -31°		190/195°, -26/-35°		Hydra/ Centaurus	D
40	Dec. 12-20	7	230°, -3°		228/231°, -2/-3°		Libra/ Serpens	B

* Dates in this table are based on N.Z.S.T. Dates shown in brackets represent activity present immediately before or after a non-operative period. Such showers may have started or terminated earlier or later respectively than the date shown.

northern hemisphere daylight showers. They occur in the June period when the aerial was not being turned daily, and hence it can only be shown that the known radiants do in fact give rise to envelopes at the time expected. By this means, however, these two showers have definitely been confirmed, and part of the morning activity seen in Figure 1 is due to their presence. Shower VII was not clearly found as it also occurred in the June period of aerial non-rotation. It will be shown in Section V that it is possible to include numbers VIII and XIII in a more general radiant scheme. Consequently number IV, referring to an April-May γ-Capricornid group is the only visual shower for which no evidence at all has been found.

(a) Night Radar Showers

It is remarkable that at the time of the very strong northern hemisphere June daytime activity there is also very strong southern hemisphere night activity (Fig. 1). In July all the strong southern hemisphere visual showers reported by McIntosh (1934, 1935, 1936, 1940) have been found, and no new activity is apparent in this month. No other particularly strong night activity has been found, but there is considerable confused activity during the earlier part of December. No visual night activity at all has been reported during December, January, February, or March. Long hours of daylight would tend to engender this result for visual observations in these months, so it is of interest to find from the radar observations that the latter half of December, February, and March do in fact appear to be largely inactive. (January has not yet been sufficiently investigated by radar to allow an estimate of activity to be given.)

(b) Daylight Radar Showers

No previous observations exist for southern hemisphere daylight meteor activity.* A weak grade D shower is possible in February. A new shower

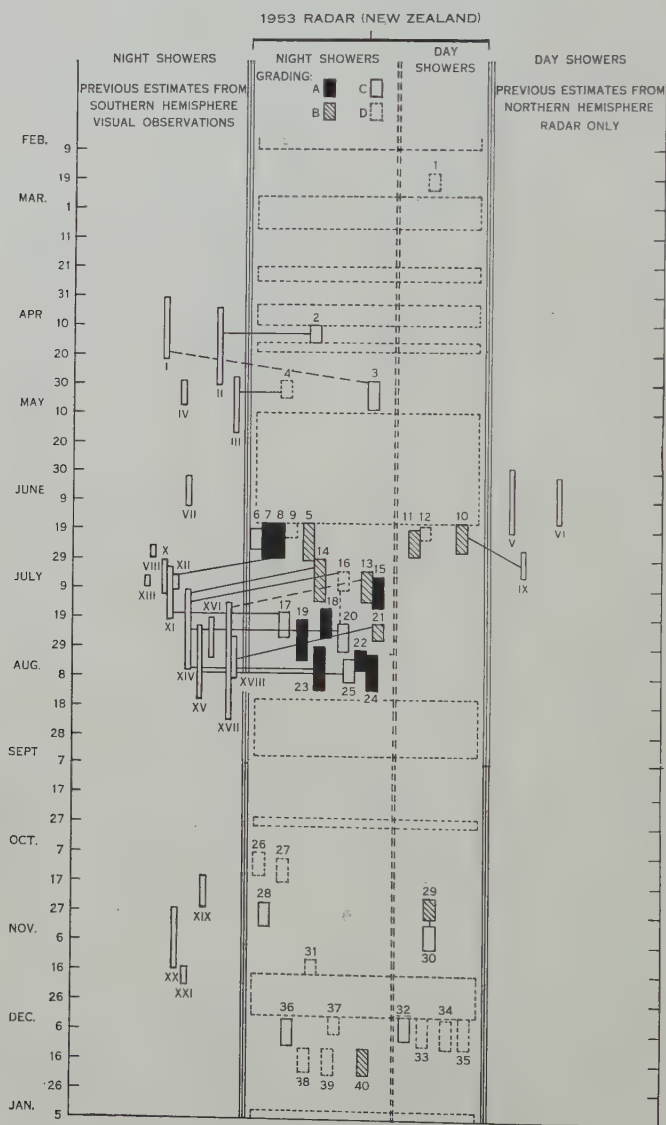


Fig. 3.—Comparison of visual and radar estimates of southern hemisphere meteor shower activity. (The dotted rectangles covering the whole of the night and day radar columns indicate periods when the equipment was inoperative.)

* Note added in Proof.—Subsequent to the submission of this paper for publication a paper has been published by Weiss (1955) in which six showers have been investigated in some detail at Adelaide by radar methods. The results are in general agreement with those in the present paper. Weiss mentions daytime activity on five scattered days in July but finds no daytime showers.

from the direction of Eridanus is present in late June, and there appears to be a strong and quite enduring shower from the direction Virgo over the period October 22–November 7. Early December activity is again confused, but some weaker radiants appear to be present. No significant daylight activity has been found over the other months.

V. THE SPATIAL DISTRIBUTION OF METEOR RADIANTS

Southern hemisphere night meteor activity is very strong between mid June and mid August, with a number of radiants clearly defined but overlapping in date. When the activity is grouped as in Figure 4 a marked drift in the origin of the meteor activity over the two-month period becomes apparent.

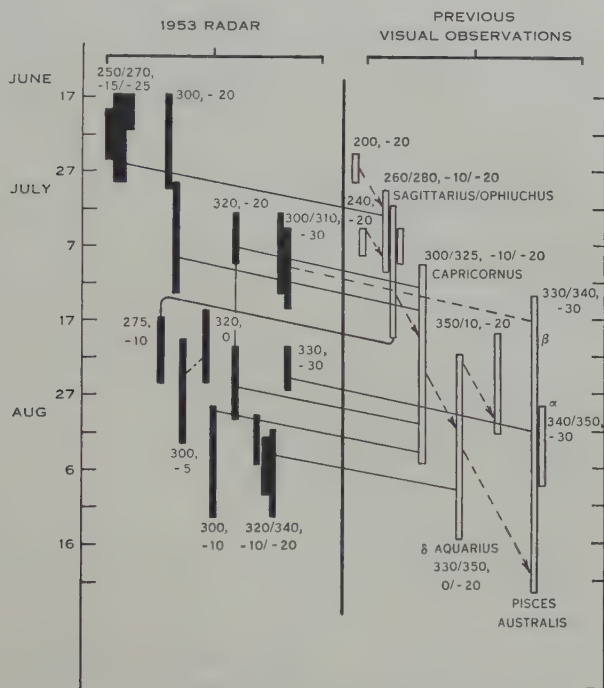


Fig. 4.—Interrelation of major night meteor showers in the period June–August.

The visual observations (based on Table 1) commence with possible activity from the region $(200^\circ, -20^\circ)$ of the sky. The Right Ascension drifts with date through the region 240° , and then $260/280^\circ$. At the same time there is a progressive downward drift in declination, as shown by the connexion of showers with dotted arrows. The Capricornids, δ -Aquadrids, and Pisces Australids are all very strong showers which fit into this progression, concluding about mid August with coordinates $(340/350^\circ, -30^\circ)$.*

* *Note added in Proof.*—The drift within the δ -Aquadrid shower radiant found by Weiss (1955) conforms with the general drift of all showers in the July–August period as found in the present paper.

The same tendency is evident from an examination of the radar results. Every group except one can be paired with a visual report, and the coordinates and date of the one isolated group are not greatly divergent from the general trend.

This passage of activity across the sky with date is similar to the northern hemisphere daylight activity (Almond, Bullough, and Hawkins 1952; Hawkins and Almond 1952*a*, 1952*b*; Bullough 1954) and led to the construction of Figure 5, which is a plot of the divergence of mean shower radiants from the ecliptic. The two dotted lines represent a divergence of 10° from the plane of the ecliptic, and both northern and southern meteor radiants have been included,

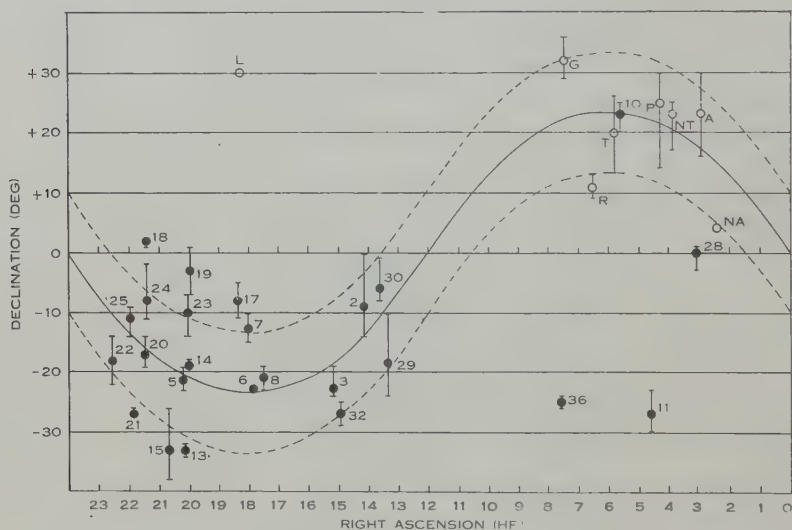


Fig. 5.—Meteor radiants plotted in relation to the plane of the ecliptic.
A, Arietid; G, Geminid; L, Lyrid; NA, night Arietid; NT, night Taurid;
P, ζ -Perseid; R, Orionid; T, β -Taurid.

excepting grade D showers of Table 2. Three northern hemisphere showers, namely the Ursids, Quadrantids, and Perseids, have not been shown on Figure 5 as they are of very high northern declination. The present radar apparatus as mentioned earlier, is not suitable for the detection of high southern radiants so similar high declination southern showers may yet be found to exist. It is clearly apparent, however, from Figure 5, that a high proportion of the major meteor shower radiants lie close to the plane of the ecliptic.

VI. ACKNOWLEDGMENTS

This paper, covering the results of an initial survey of southern hemisphere meteor activity, is part of a programme of radar-meteor observations which has been made possible by grants received from the Nuffield Foundation, the University of New Zealand, and the New Zealand Department of Scientific and Industrial Research.

VII. REFERENCES

- ALMOND, M., BULLOUGH, K., and HAWKINS, G. S. (1952).—*Jodrell Bank Ann.* **1**: 13.
ASPINALL, A., CLEGG, J. A., and HAWKINS, G. S. (1951).—*Phil. Mag.* **42**: 504.
BULLOUGH, K. (1954).—*Jodrell Bank Ann.* **1**: 68.
CLEGG, J. A. (1948).—*Phil. Mag.* **39**: 577.
HAWKINS, G. S., and ALMOND, M. (1952a).—*Jodrell Bank Ann.* **1**: 2.
HAWKINS, G. S., and ALMOND, M. (1952b).—*Mon. Not. R. Astr. Soc.* **112**: 219.
HOFFMEISTER, C. (1948).—"Meteorströme." (Weimar.)
LINDBLAD, B. (1952).—Chalmers Tek. Högsk. Handl. No. 129.
MCINTOSH, R. A. (1934).—*Trans. Roy. Soc. N.Z.* **63**: 433.
MCINTOSH, R. A. (1935).—*Mon. Not. R. Astr. Soc.* **95**: 709.
MCINTOSH, R. A. (1936).—*Trans. Roy. Soc. N.Z.* **66**: 60.
MCINTOSH, R. A. (1940).—*Trans. Roy. Soc. N.Z.* **69**: 392.
MCKINLEY, D. W. R. (1954).—*Astrophys. J.* **119**: 519.
PRENTICE, J. P. M. (1955).—*Handb. Brit. Astr. Ass.* **1955**: 42.
WEISS, A. A. (1955).—*Aust. J. Phys.* **8**: 148.

THE DISTRIBUTION OF THE RATIO OF TWO QUADRATIC FORMS*

By G. S. WATSON†

[Manuscript received March 31, 1955]

Summary

The exact probability distribution of the ratio of two quadratic forms is given in the case where the quadratic forms and the multivariate normal distribution of the variables is such that the ratio is a ratio of linear functions of gamma variables of integral order.

I. INTRODUCTION

The distribution of the ratio of two quadratic forms in normal variables is required for a variety of problems, especially in the analysis of time series. In general it is required for non-null distributions of serial correlation coefficients. Dixon (1944) proposed the problem of finding the distribution of

$$r = \frac{\sum_{i=1}^N (x_i - 2x_{i+1} + x_{i+2})^2}{\sum_{i=1}^N (x_i - x_{i+1})^2}, \dots\dots\dots (1)$$

with $x_{N+1}=x_1$, $x_{N+2}=x_2$, when x_1, \dots, x_N are N.I.D. (0,1). Dixon found the mean and variance of the smoothed distribution of this ratio.

Gurland (1953) has shown how to express the distribution of

$$r = \frac{\mathbf{x}'\mathbf{A}\mathbf{x}}{\mathbf{x}'\mathbf{B}\mathbf{x}}, \dots\dots\dots (2)$$

where \mathbf{x} is a vector of N.I.D. (0,1) variables and \mathbf{A} and \mathbf{B} are real symmetric matrices, \mathbf{B} positive definite, by a Laguerrian expansion. Pitman and Robbins (1949) give a different series solution for the same problem. In many cases of interest, however, these series converge very slowly and are therefore of no practical value. The Pitman and Robbins method was applied by the present writer to the distribution of von Neumann's ratio and the series found by using the EDZAC, Cambridge Mathematical Laboratory; the terms of the series did not begin to decrease until the 37th term! Less elegant methods (which give no estimate of error) are thus still necessary in many problems. It is therefore worth while to have an exact distribution, which is available in certain cases, to provide a check.

* This paper was written when the author was Research Officer, Department of Applied Economics, University of Cambridge, and formed a part of a thesis issued by the Institute of Statistics, University of North Carolina, Mimeograph Series.No. 49, 1951.

† Australian National University, Canberra, A.C.T.

II. AN EXACT DISTRIBUTION

We suppose that the matrices **A** and **B** of (2) are commutative so that they may be diagonalized by the same orthogonal transformation and, furthermore, that their latent roots are equal in pairs. Then we may write

$$r = \frac{\sum_1^n \lambda_j w_j}{\sum_1^n \mu_j w_j} = \frac{l}{m}, \text{ say, } \dots\dots\dots (3)$$

where w_j ($j=1, \dots, n$) are independent $\gamma(1)$ variables, and where λ_j, μ_j ($j=1, \dots, n$) satisfy the conditions

$$\left. \begin{aligned} J_{jj'} &= \begin{vmatrix} \lambda_j & \mu_j \\ \lambda_{j'} & \mu_{j'} \end{vmatrix} > 0 & (\text{all } j < j'), \\ \mu_j &> 0 & (\text{all } j), \\ \begin{vmatrix} 1 & \lambda_k & \mu_k \\ 1 & \lambda_j & \mu_j \\ 1 & \lambda_{j'} & \mu_{j'} \end{vmatrix} &\neq 0 & (\text{all } k \neq j \neq j'). \end{aligned} \right\} \dots\dots\dots (4)$$

It may be verified that the ratio (1) may be put into a form (3) satisfying all these conditions if N is odd. The same is true of the circular lag-1 serial correlation with a mean correction when the population is a circular second order autoregressive process.

Before we can proceed we need the following

Lemma

If, for any real numbers λ_k, μ_k ($k=1, \dots, n$),

$$\begin{vmatrix} 1 & \lambda_k & \mu_k \\ 1 & \lambda_j & \mu_j \\ 1 & \lambda_{j'} & \mu_{j'} \end{vmatrix} \neq 0 \quad (\text{all unequal } k, j, j'),$$

then

$$\frac{1}{\prod_{k=1}^n (1 - \lambda_k u - \mu_k v)} = \sum \sum_{j < j'} \frac{\begin{vmatrix} \lambda_j & \mu_j \\ \lambda_{j'} & \mu_{j'} \end{vmatrix}^{n-2}}{\prod_{k \neq j, j'} \begin{vmatrix} 1 & \lambda_k & \mu_k \\ 1 & \lambda_j & \mu_j \\ 1 & \lambda_{j'} & \mu_{j'} \end{vmatrix}} \frac{1}{(1 - \lambda_j u - \mu_j v)(1 - \lambda_{j'} u - \mu_{j'} v)}.$$

Proof

For $n=3$, the lemma may be simply proved by considering the identity $1 = \sum_{k=1}^3 A_k (1 - \lambda_k u - \mu_k v)$. Successive applications of the result for $n=3$ then give the general result.

The joint characteristic function of l and m , $\varphi(u, v)$, say, is given by

$$\varphi(u, v) = \prod_{j=1}^n (1 - iu\lambda_j - iv\mu_j)^{-1} \\ = \sum_{j < j'} \frac{\left| \begin{array}{cc} \lambda_j & \mu_j \\ \lambda_{j'} & \mu_{j'} \end{array} \right|^{n-2}}{\prod_{k \neq j, j'} \left| \begin{array}{ccc} 1 & \lambda_k & \mu_k \\ 1 & \lambda_j & \mu_j \\ 1 & \lambda_{j'} & \mu_{j'} \end{array} \right|} \frac{1}{(1 - i\lambda_j - i\mu_{j'}v)(1 - i\lambda_{j'}u - i\mu_jv)}, \quad \dots (5)$$

by the lemma. Thus the joint probability density of l and m is given by

$$f(l, m) = \frac{1}{(2\pi i)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \varphi(u, v) e^{-iul - ivm} du dv, \quad \dots (6)$$

where the integration is to be performed separately on every term of (5). But

$$\frac{1}{(2\pi i)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\exp(-iul - ivm)}{(1 - i\mu\lambda_j - iv\mu_j)(1 - iu\lambda_{j'} - iv\mu_{j'})} du dv,$$

is the joint probability density of

$$l_{jj'} = \lambda_j w_j + \lambda_{j'} w_{j'}, \quad m_{jj'} = \mu_j w_j + \mu_{j'} w_{j'},$$

if

$$\frac{\lambda_j}{\mu_j} \geq \frac{l}{m} \geq \frac{\lambda_{j'}}{\mu_{j'}},$$

or zero otherwise. This density may be found *ab initio* and is given by

$$\frac{e^{-1}}{J_{jj'}} \exp \left\{ \frac{1}{J_{jj'}} \left| \begin{array}{ccc} 1 & l_{jj'} & m_{jj'} \\ 1 & \lambda_j & \mu_j \\ 1 & \lambda_{j'} & \mu_{j'} \end{array} \right| \right\}, \quad \dots (7)$$

since $J_{jj'} > 0$ by (4). Thus,

$$\frac{\lambda_i}{\mu_i} \geq \frac{l}{m} \geq \frac{\lambda_{i+1}}{\mu_{i+1}}.$$

$$f(l, m) = \sum_{j=1}^i \sum_{j'=i+1}^n \frac{J_{jj'}^{n-3} e^{-1}}{\prod_{k \neq j, j'} \left| \begin{array}{ccc} 1 & \lambda_k & \mu_k \\ 1 & \lambda_j & \mu_j \\ 1 & \lambda_{j'} & \mu_{j'} \end{array} \right|} \exp \left\{ \frac{1}{J_{jj'}} \left| \begin{array}{ccc} 1 & l & m \\ 1 & \lambda_j & \mu_j \\ 1 & \lambda_{j'} & \mu_{j'} \end{array} \right| \right\}. \quad \dots (8)$$

Now

$$\frac{1}{J_{jj'}} \left| \begin{array}{ccc} 1 & l & m \\ 1 & \lambda_j & \mu_j \\ 1 & \lambda_{j'} & \mu_{j'} \end{array} \right| = 1 - \frac{1}{J_{jj'}} \{ (\mu_{j'} - \mu_j)l - (\lambda_{j'} - \lambda_j)m \},$$

so the joint density of the variables $r=l/m$ and m is

$$f(r, m) = \sum_{j=1}^i \sum_{j'=i+1}^n \frac{J_{jj'}^{n-3} m}{\prod_{k \neq j, j'} \begin{vmatrix} 1 & \lambda_k & \mu_k \\ 1 & \lambda_j & \mu_j \\ 1 & \lambda_{j'} & \mu_{j'} \end{vmatrix}} \times \exp \left\{ -\frac{m}{J_{jj'}} [(\mu_{j'} - \mu_j)r - (\lambda_{j'} - \lambda_j)] \right\}, \quad \dots \quad (9)$$

for $\lambda_i/\mu_i \geq r \geq \lambda_{i+1}/\mu_{i+1}$. For fixed r , m may run between zero and infinity. Integrating out m , the density of r is found to be

$$f(r) = \sum_{j=1}^i \sum_{j'=i+1}^n \frac{J_{jj'}^{n-1}}{\prod_{k \neq j, j'} \begin{vmatrix} 1 & \lambda_k & \mu_k \\ 1 & \lambda_j & \mu_j \\ 1 & \lambda_{j'} & \mu_{j'} \end{vmatrix}} \frac{1}{[(\mu_{j'} - \mu_j)r - (\lambda_{j'} - \lambda_j)]^2}, \quad \dots \quad (10)$$

for $\lambda_i/\mu_i \geq r \geq \lambda_{i+1}/\mu_{i+1}$. Integrating with respect to r , the distribution function of r is given by

$$P(r \geq r_0) = \sum_{i=1}^i \sum_{j=1}^i \sum_{j'=i+1}^n \frac{J_{jj'}^{n-1}}{\prod_{k \neq j, j'} \begin{vmatrix} 1 & \lambda_k & \mu_k \\ 1 & \lambda_j & \mu_j \\ 1 & \lambda_{j'} & \mu_{j'} \end{vmatrix}} \frac{J_{i, i-1}}{\begin{vmatrix} 0 & \lambda_{i-1} & \mu_{i-1} \\ 1 & \lambda_j & \mu_j \\ 1 & \lambda_{j'} & \mu_{j'} \end{vmatrix}} \frac{\begin{vmatrix} 0 & \lambda_i & \mu_i \\ 1 & \lambda_j & \mu_j \\ 1 & \lambda_{j'} & \mu_{j'} \end{vmatrix}}{\begin{vmatrix} 0 & \lambda_{i-1} & \mu_{i-1} \\ 1 & \lambda_j & \mu_j \\ 1 & \lambda_{j'} & \mu_{j'} \end{vmatrix}} + \sum_{j+1}^i \sum_{j'=i+1}^n \frac{J_{jj'}^{n-1}}{\prod_{k \neq j, j'} \begin{vmatrix} 1 & \lambda_k & \mu_k \\ 1 & \lambda_j & \mu_j \\ 1 & \lambda_{j'} & \mu_{j'} \end{vmatrix}} \frac{(\lambda_{i_0} - r_0 \mu_{i_0})}{\begin{vmatrix} 0 & \lambda_i & \mu_i \\ 1 & \lambda_j & \mu_j \\ 1 & \lambda_{j'} & \mu_{j'} \end{vmatrix}} \frac{\begin{vmatrix} 0 & r & 1 \\ 1 & \lambda_j & \mu_j \\ 1 & \lambda_{j'} & \mu_{j'} \end{vmatrix}}{\begin{vmatrix} 0 & \lambda_i & \mu_i \\ 1 & \lambda_j & \mu_j \\ 1 & \lambda_{j'} & \mu_{j'} \end{vmatrix}}, \quad \dots \quad (11)$$

for

$$\frac{\lambda_{i_0}}{\mu_{i_0}} \geq r_0 \geq \frac{\lambda_{i_0+1}}{\mu_{i_0+1}}.$$

III. AN ALTERNATIVE FORM

The expression (11) would be difficult to calculate, even for small n . However, the use of the ratio device leads to a much simpler result. Box (1954) has given a general theorem on the distribution of ratios of quadratic forms of the type we are dealing with. Instead of stating his theorem, it is preferable to derive the distribution directly by his methods.

Writing $v_j = R\mu_j - \lambda_j$, we have

$$P(r \leq R) = P(\sum (\lambda_j - R\mu_j)w_j \leq 0) \\ = P(\sum v_j w_j \geq 0). \quad \dots \quad (12)$$

To evaluate (12), we notice that the characteristic function of $\Sigma v_j w_j$, $\varphi(t)$, admits the partial fraction expansion

$$\varphi(t) = \sum_{j=1}^n \frac{v_j^{n-1}}{\prod_{\substack{j'=1 \\ (j' \neq j)}}^n (v_j - v_{j'})} \frac{1}{(1 - i v_j t)} \quad (v_j \text{'s distinct}). \quad \dots \quad (13)$$

Thus the density of $\Sigma v_j w_j$, $f(x)$, say, is given by

$$\begin{aligned} f(x) &= \frac{1}{2\pi i} \int_{-\infty}^{\infty} e^{-ixt} \varphi(t) dt \\ &= \sum_{j=k}^n \frac{v_j^{n-2}}{\prod (v_j - v_{j'})} \exp\left(-\frac{x}{v_j}\right), \quad \dots \dots \dots (14) \end{aligned}$$

where k is defined by $\lambda_k/\mu_k \leq R \leq \lambda_{k-1}/\mu_{k-1}$. Thus

$$\begin{aligned} P(r \leq R) &= \int_0^{\infty} f(x) dx \\ &= \sum_{j=k}^n \frac{v_j^{n-1}}{\prod (v_j - v_{j'})} \\ &= \sum_{j=k}^n \frac{(R\mu_j - \lambda_j)^{n-1}}{\prod_{j'} [(\mu_j - \mu_{j'})R - (\lambda_j - \lambda_{j'})]}. \quad \dots \dots \dots (15) \end{aligned}$$

This is the required distribution. It is important to note that, had we sought by these methods an expression for $P(r \geq R)$ we would have found

$$P(r \geq R) = \sum_{j=1}^k \frac{(\lambda_j - R\mu_j)^{n-1}}{\prod_{j'} [(\lambda_j - \lambda_{j'}) - R(\mu_j - \mu_{j'})]}, \quad \dots \dots \dots (16)$$

where k would be defined by $\lambda_k/\mu_k \geq R \geq \lambda_{k+1}/\mu_{k+1}$.

Since $P(r \leq R) + P(r \geq R) = 1$, we can use for calculation the formula with the shorter summation. In applications of Box's theorem, two formulae will always be available and the shorter will naturally be used.

It will be noted that, if all the μ_j equal unity, the formula (16) reduces to that given by Anderson (1942) for the distribution of the first circular serial correlation coefficient, as it should.

The expressions (15) and (16) are not as convenient as (10) for the calculation of the exact moments of r . In practice, approximate moments could be found by using the easily determined moments of l and m .

IV. ACKNOWLEDGMENT

The author is indebted to the referee for the reference to Box's work.

V. REFERENCES

- ANDERSON, R. L. (1942).—Distribution of the serial correlation coefficient. *Ann. Math. Statist.* **13**: 1–13.
- Box, G. E. P. (1954).—Some theorems on quadratic forms . . . *Ann. Math. Statist.* **25**: 290–302.
- DIXON, W. J. (1944).—Further contributions to the problem of serial correlation. *Ann. Math. Statist.* **15**: 119–44.
- GURLAND, J. (1953).—Distribution of quadratic forms and ratios of quadratic forms. *Ann. Math. Statist.* **24**: 416–27.
- PITMAN, E. J. G., and ROBBINS, H. (1949).—Application of the method of mixtures to quadratic forms in normal variates. *Ann. Math. Statist.* **20**: 552–60.

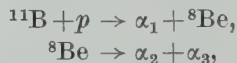
SHORT COMMUNICATIONS

THE DESIGN OF A MULTIPLE-WIRE PROPORTIONAL COUNTER AND ITS USE IN THE STUDY OF THE α -PARTICLES FROM THE REACTION ${}^7\text{Li}+p^\dagger$

By A. C. RIVIERE ‡ and P. B. TREACY †

Introduction

This counter was designed originally for the study of a two stage break-up process such as



where correlations are sought between α_1 and α_2 as discussed by Treacy (1955). In this case the α -particles α_2 and α_3 are emitted into a cone of semi-angle 10° and could readily be accepted by the counter. The design was not specific to this experiment but enables α -particles of energies up to 8 MeV, falling within a cone of semi-angle 60° to be detected with electron collection times of less than $1\ \mu\text{sec}$.

A proportional counter was chosen but, since the particle ranges must be contained wholly within the sensitive region of the counter, it is difficult to achieve fast electron collection in a single-wire counter with an E.H.T. supply of the order of 1 kV. This difficulty can be avoided by subdividing the sensitive region into a number of separate counters, each with its anode and cathode. In such a system the speed of collection can be increased considerably without changing anode wire diameter or E.H.T. voltage.

Construction

Figure 1 shows a schematic cross section of the counter. The sensitive region is within a cube of side 8 in. containing 16 independent anode-cathode systems. Each anode wire A is located at the centre of a square array of eight thicker wires C all earthed and which form the cathode. The dotted lines in Figure 1 indicate the cone of semi-angle 60° within which particles from a target T at the cone apex can be detected and which has no line of axial symmetry. Thus detection occurs with equal speed and efficiency for all directions within this cone.

The anode wires are of 0.0035 in. diameter beryllium-copper and are attached at their ends to two ebonite sheets mounted by screws on a brass frame F . This frame has large holes H drilled through it to expose the ebonite E

† Manuscript received February 10, 1955.

‡ Research School of Physical Sciences, Australian National University, Canberra.

around each anode wire and thus allow the necessary insulation. The common anode lead emerges through a spring-loaded contact to a "Kovar"-glass seal and is arranged so that the top flange of the vacuum box, with frame attached, can be lifted out of the cylinder. The cathode wires are of 0.010 in. diameter copper and are attached directly to the frame.

The outer vacuum box comprises a 12 in. diameter cylinder (rolled from 0.25 in. brass sheet) with rubber vacuum seals. It may be evacuated through a glass tube held by a semi-flexible O-ring seal. There is provision elsewhere for two similar glass seals, so that the counter gas may be circulated over sodium metal for purification. A thin mica window, supported on a brass grid, is shown in Figure 1 attached to the target box. The holes of the grid are countersunk at the back to allow wide angle penetration of the counter.

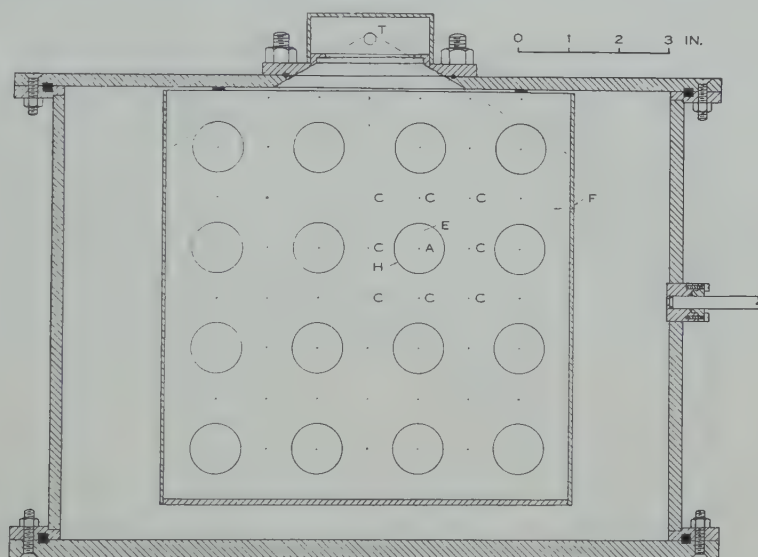


Fig. 1.—Schematic view of the counter in cross section.

Performance

The counter is used in conjunction with a linear amplifier having rise and clipping times of 0.16 and 1.6 μsec respectively. With the counter filled to 38 cm Hg pressure with argon (99.8 per cent. pure) and operated at an E.H.T. voltage of 720 V, a 5.3 MeV α -particle, after traversing the 7 mm mica window, produces a pulse equivalent to 7.65×10^5 ion pairs. This represents a gas amplification factor of about 4.7 times with a noise equivalent at the input of 2.7×10^4 ion pairs. The noise in the absence of the counter capacity, 56 $\mu\mu\text{F}$, is effectively 7.8×10^3 ion pairs, the reduction being ascribed to shot effect.

In practice the gas gain can be increased up to the Geiger region; moreover, the performance of the counter is quite conventional. As is shown below, in discussing Figure 2, the counter is proportional, having a pulse spread for 9.0 MeV α -particles of approximately ± 6 per cent. at half height: this spread is attributed to the grid-like construction and "square" geometry of each

cathode, which results in variation of electric field in different directions out from the anode wire, and also to small non-uniformities of wire diameter.

Under the conditions quoted above, the pulse rise time due to positive ion motion was measured as $2.4 \mu\text{sec}$, using the method given by Gillespie (1953). This represents a constant delay in pulse rise which is limited by the size of the anode wire and could be reduced if desired. The electron collection time, which is subject to fluctuations, is estimated to be less than $1 \mu\text{sec}$. Good linearity was obtained in the experiment referred to in the introduction, with a coincidence resolving time of $1.2 \mu\text{sec}$.

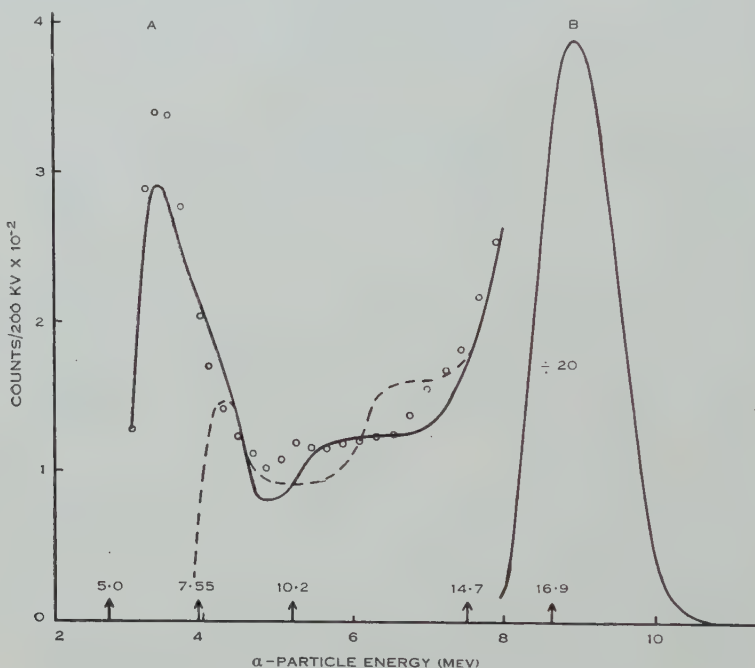
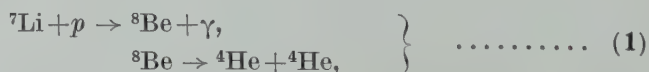


Fig. 2.— α -Particle spectra from ${}^7\text{Li}+p$. Full curve, spectrum taken at proton energy of 0.47 MeV ; circles, mean of spectra at proton energies of 0.40 and 0.54 MeV ; broken curve, spectrum of ${}^{210}\text{Po}$ α -particles. All curves are normalized to 5×10^4 counts in the main group B.

The α -Particles from the Reaction ${}^7\text{Li}+p$

The behaviour of the counter can be illustrated by results obtained on the spectrum of α -particles produced by bombarding lithium with protons of energies in the neighbourhood of the 440 keV resonance. This experiment, although of interest in the study of ${}^8\text{Be}$, has not previously been carried out in any detail (Ajzenberg and Lauritsen 1955). α -Particles can be produced by the following processes :



In reaction (1) the various ^8Be levels give rise to α -particles of different energies. Burcham and Freeman (1950) have analysed the α -particles of energy up to 2.4 MeV from reaction (1) and found that the ground and first excited states of ^8Be are formed as intermediate stages. In reaction (2) a direct process occurs giving a single group of α -particles of energy 8.98 MeV (Strait *et al.* 1951). α -Particles of energies between these limits would be expected corresponding to other ^8Be levels. These levels would be formed by γ -ray emission from the 17.63 MeV state formed at the 440 keV resonance (Inall 1954) unless such photon emission is of high multipolarity resulting in a relatively low intensity.

To examine these possibilities an experiment was carried out as follows. The counter was set at 90° to the proton beam of the Canberra 1.2 MeV accelerator, the beam being collimated by a $\frac{3}{16}$ in. diameter aperture and allowed to strike a 50 keV thick, solid-backed target of lithium aluminium hydride. α -Particles passed through an aperture subtending $\pm 10^\circ$ at the target and entered the counter through a mica window of 7 mm air equivalent. Each pulse from the counter was displayed on an oscillograph and recorded photographically. The α -particle spectra detected at 0.40, 0.47, and 0.54 MeV proton energies are represented in Figure 2. The run at 0.47 MeV proton energy is drawn as a full curve and, for comparison, only the points of the mean of the upper- and lower-energy runs are plotted. It is evident that no resonant effects are present. The positions of known ^8Be levels are indicated by arrows labelled with the appropriate excitation energies. The resolved peaks *A* and *B* are due to $^{18}\text{O}(p\alpha)^{15}\text{N}$ and reaction (2) respectively. All results are normalized to 5×10^4 counts in the main group *B*.

Apart from these a background is present which is due to the counter. This can be seen from the dotted curve in Figure 2 which is the spectrum obtained for a ^{210}Po α -particle source where the counter gas pressure has been reduced to 16 cm Hg. At this pressure the range of ^{210}Po α -particles in the counter is equal to that of the α -particles from reaction (2) under the previous conditions. A range-energy relation for argon was derived from data given for air by Bethe (1950), the non-linearity of which accounts for the difference in shape between the dotted and full curves.

From the known target thickness, data on total α -particle and γ -ray yields (Boyle and Inall, personal communication 1954) and by using the yield curve from Heydenburg *et al.* (1948) it can be shown from the results that, if ^8Be levels between 7 and 15 MeV are excited, their intensity must be less than 3 per cent. of the intensity of the transitions to the ground state. As the transition to the ^8Be level at 7.5 MeV is only of relative intensity 1 per cent. (Inall 1954), this would not have been detected above the background in the present experiment. Transitions by reaction (1) to levels higher than 15 MeV (Wilkins and Goward 1953) would here be masked by the α -particles produced by reaction (2).

This work was done as part of the research programme of this laboratory under Professor E. W. Titterton. The authors are indebted to Mr. R. G. Hawkins for much technical assistance.

References

- AJZENBERG, F., and LAURITSEN, T. (1955).—*Rev. Mod. Phys.* **27** : 77.
 BETHE, H. A. (1950).—*Rev. Mod. Phys.* **22** : 213.
 BURCHAM, W. E., and FREEMAN, J. M. (1950).—*Phil. Mag.* **41** : 921.
 GILLESPIE, A. B. (1953).—"Signal, Noise and Resolution in Nuclear Counter Amplifiers."
 (Pergamon Press Ltd. : London.)
 HEYDENBURG, N. P., HUDSON, C. M., INGLIS, D. R., and WHITEHEAD, W. D. (1948).—*Phys. Rev.*
74 : 405.
 INALL, E. K. (1954).—*Phil. Mag.* **45** : 768.
 STRAIT, E. N., VAN PATTTER, D. M., SPERDUTO, A., and BUECHNER, W. W. (1951).—*Phys. Rev.*
81 : 747.
 TREACY, P. B. (1955).—*Proc. Phys. Soc. Lond.* A **68** : 204.
 WILKINS, J. J., and GOWARD, F. K. (1953).—*Proc. Phys. Soc. Lond.* A **66** : 661.

ON STARTING ROUTINES FOR THE C.S.I.R.O. MARK I COMPUTER*

By G. W. HILL† and T. PEARCEY†

Introduction

High speed computers have a normal mode of operation whereby operations are specified by "commands" from stored programmes. This mode cannot be used for starting a cleared computer since there are no stored commands to perform the organized process of assembling commands from the input medium. Additional equipment is therefore required in order to provide another mode of operation for inserting the first, or primary, routine.

Primary Routine

The C.S.I.R.O. Mark I computer can insert a primary routine from stepping switches, as is described in previous papers (Pearcey and Hill 1953*a*, 1953*b*) from which the notation in this article is adopted. Other non-normal modes of operation are included for purposes of testing the computer. One of these modes is used to insert the primary routine, listed in Table 1, from punched tape into the store. The special equipment required for the previous primary routine is therefore redundant.

For testing purposes, commands can be taken from the switchboard of the console, instead of from the store, by special modes of operation controlled from the switchboard. A switch denoted as " $S \& N_1$ to INT " enables performance of the logical sum of sequence register contents (increasing serially) and the command on the switch-set register N_1 . If this command is " $(I) \rightarrow M$ " then this mode of operation reads rows of punched tape from the tape reader into serial locations in the store. In this manner the primary routine listed in Table 1 is inserted into the cleared computer.

* Manuscript received May 18, 1955.

† Division of Radiophysics, C.S.I.R.O., University Grounds, Sydney.

Since the input equipment of the C.S.I.R.O. Mark I computer reads 10-hole tape rows into digit positions p_1 - p_{10} and the X and Y holes into digit positions p_{19} and p_{20} , respectively, the primary routine must not require digits in the positions p_{11} - p_{18} . This adds a fourth requirement to the three requirements of any primary routine. Thus the primary routine must:

(1) assemble 20-digit commands from two 10-digit tape rows according to the convention that the address-digits row (if any) is followed by an X -punched operation code row;

(2) insert assembled commands in serial locations of the store;

(3) cause Y -punched "control designations" to shift control out of the primary routine for performance of special operations;

(4) be composed of digits in the p_1 - p_{10} and p_{19} , p_{20} positions only.

The primary routine set out in Table 1 satisfies all these requirements.

TABLE 1
PRIMARY ROUTINE

Location	Command		Operation
0	$(0) \xrightarrow{c} S^*$	[4]	Command itself contains 3 digits, shifts control to 4
1	$(C) \xrightarrow{+} K$		Address in C is added to the following command
2	$c(A) \rightarrow 0$		Transfers command in A to store location specified by (C)
3	$p_{11} \xrightarrow{+} C$		Adds unity to the address in C for serial storage
4	$(Z) \rightarrow B$		Sets $(B)=0$; indicating "non- X " row for command 15
5	$(D_0) \rightarrow H_l$		Sets 10-digit input into H
6	$(D_0) \xrightarrow{+} D_0$		D_0 contains p_{20} only if X -hole just recently read
7	$s(D_0) \xrightarrow{c} S$		Tests whether recently read row is X -punched or not
8	$(0) \xrightarrow{c} S$	[12]	Jumps three commands to 12, to add non- X punched row to upper half
9	$p_{11} \rightarrow B$		Sets $(B)=p_{11}$; indicating " X -row" for command 15
10	$(H_l) \xrightarrow{+} A$		Adds operation code to p_1 - p_{10} positions to assemble command
11	$p_1 \xrightarrow{c} S$		Jump next command
12	$(H_u) \xrightarrow{+} A$		Adds address code to p_{11} - p_{20} positions to assemble command
13	$(I) \rightarrow D_0$		Reads tape row into D_0 with p_{19} for X , p_{20} for Y
14	$s(D_0) \xrightarrow{c} S$		If Y -punched row, jump to command 16
15	$(B) \rightarrow S$	[0, 1]	Shift to 1 or 0 if previous row X -punched or not (resp.)
16	$[(D_0) \rightarrow H_l]$	}	These four commands are overwritten by subsequent input after being used to set $16p_{11}$ in C by shifting $16p_1$ read from tape through H into upper half of C and then clearing D_0 and S
17	$[(H_u) \rightarrow C]$		
18	$[(D_0) \rightarrow D_0]$		
19	$[(Z) \rightarrow S]$	[0]	
	Blank tape rows		Restore normal mode of operation by hand (switch " S & N_1 to INT " off)
	16Y		Y -detection by command 14 shifts to special operation of setting the starting address in C

* Command " $(0) \xrightarrow{c} S$ " on the tape is immediately preceded by an X punch.

The first requirement is satisfied since the 10-digit row read from tape by command 13 is added as an upper half address by commands 8, 12 if commands 6, 7 indicate the row is not *X*-punched, or is added as a lower half operation code by commands 9, 10, 11 if commands 6, 7 indicate the row is *X*-punched. Serial storage of assembled commands by commands 1, 2, 3, whenever unity in *B* indicates previous row as *X*-punched, satisfies the second requirement. Command 14 detects *Y*-punched rows and shifts control out of the primary routine to command 16, thus fulfilling the third requirement.

To satisfy the fourth requirement, this primary routine uses commands with zero addresses and holds necessary addresses in registers. Sequence shifts, which normally imply addresses, are replaced by the "sequence jump" command " $(0) \xrightarrow{c} S$ ", which jumps three commands for the three digits in the command itself. Register *B* is used to hold addresses 1 or 0 corresponding to *X*-hole or blank on the row previously read. The "storage address" in register *C* is set initially by commands 16, 17, 18, 19 as soon as the routine reads the 16*Y* row and this address is combined with command 2 by use of the " $+K$ " facility to store subsequently assembled commands in serial locations from 16 onwards.

Insertion of the primary routine requires a sequence of switching operations on the console switchboard, which is similar to that required previously for the stepping-switch primary routine.

(1) Clear store. Clear all registers. Switch on reader motor and "tape read" switches.

(2) Set " $(I) \rightarrow T$ " on N_1 . Set " $S \& N_1$ to *INT*".

(3) Start computer. Computer will stop after initial "*X*" is read in.

(4) Set " $(I) \rightarrow M$ " on N_1 . Clear sequence register and restart.

(5) Just before "16*Y*" is read, switch " $S \& N_1$ to *INT*" off, without stopping the machine.

(6) Further commands punched in the normal convention will be stored serially from location 16 onwards.

Control Routine

In most cases, as soon as the primary routine is inserted, it is used to insert a "control routine" from tape. This control routine provides special operations useful in assembling elaborate programmes by suspending command assembly and serial storage while either

(1) the command just assembled is performed instead of being stored,

(2) the "storage address", n_1 , is replaced by n ,

(3) the current storage address, n_r , is placed in location n of a parameter block, or

(4) the parameter in location n of the parameter block is added to the next command assembled,

where n is the address number recently read from tape.

The control routine listed in Table 2 is inserted in store locations 16–28 following the primary routine. When a *Y*-punched "control designation" is detected by command 14, control is shifted to command 16 of the control routine.

The number, x , punched in the "control designation" row is used by commands 17, 18 to shift control to command $19+x$.

For $x \leq 3$, constants are added to the address n , previously assembled in A , to form a command, which is set in location 24 and performed. Thus nT , coded as $n;0Y$, sets np_{11} in C so that serial storage will continue from address n . When the storage address is m , the control designation nS , coded as $n;2Y$, sets m in store location $28+n$, i.e. the n th parameter of the block. Subsequently, the control designation nC , coded as $n;1Y$, will reset the transfer address to that held as the n th parameter; e.g. m in location $28+n$. The designation nA , coded as $n;3Y$, sets the n th parameter, held in location $28+n$, into A so that the next command assembled will be added to this parameter and the sum stored.

TABLE 2
CONTROL ROUTINE

Location	Command	Operation
16	$(Z) \rightarrow B$	Clear B to indicate "previous column not X -punched"
17	$(D_0) \rightarrow H_l$	$\left. \begin{array}{l} \text{The "control-designation" number in the } Y\text{-punched} \\ \text{row is used to select the control operation required} \end{array} \right\}$ nT (transfer), command 24 becomes " $n \xrightarrow{\pm} C$ " $n;0Y^*$
18	$(H_u) \xrightarrow{\pm} S$	
19	$(28) \xrightarrow{\pm} A$	
20	$(27) \xrightarrow{\pm} A$	
		nC (continue), $(A) + (27) + (26) + (19) \equiv "(28+n) \rightarrow C"$ $n;1Y$
21	$(26) \xrightarrow{\pm} A$	nS (store), $(A) + (26) + (19) \equiv "(C) \rightarrow 28+n"$ $n;2Y$
22	$(19) \xrightarrow{\pm} A$	nA (add), $(A) + (19) \equiv "(28+n) \xrightarrow{\pm} A"$ $n;3Y$
23	$c(A) \rightarrow 24$	D (do), command in A set in 24 and performed $;4Y$
24	$[p_{20} \rightarrow T]$	U (unchanged), re-performs previous control operation $;5Y$
25	$13 \rightarrow S$	N (null), has no effect during insertion $;6Y$
26	$\langle 0, 0; 13, 27 \rangle$	$\left. \begin{array}{l} "(C) \rightarrow M" \text{ less } "(M) \xrightarrow{\pm} A" \\ "(M) \rightarrow C" \text{ less } "(C) \rightarrow M" \\ "(K) \rightarrow C" \text{ less } "(28) \rightarrow C" \end{array} \right\}$ Constants equal to differences between coded commands
27	$\langle 31, 31; 18, 14 \rangle$	
28	$\langle 31, 4; 26, 0 \rangle$	

* Punch codes for control designations.

A command followed by D , punched as $4Y$, is set by command 23 into 24 and performed. The last control operation performed may be repeated by using the designation U , punched as $5Y$. The null operation, N , punched as $6Y$, has no effect during insertion of commands, but incidentally is used by the routine for printing punched tape as symbolic commands to provide spaces between printed routines.

These control operations are similar to those used in association with the stepping switch primary routine. However, the number x of the control designation is now to be punched as a binary number rather than as x holes in a row. Also the operation of replacing the storage command, R , has been changed to the operation, C (continue), which restores the storage address to that held in location n of the parameter block. The main use of operation R was for replacing the storage command in order to insert programmes in the backing store instead

of the high-speed command store. This effect can be obtained by the sequence of operations :

$$\begin{aligned}c(A) &\rightarrow 2, D \\ c(A) &\rightarrow M', U.\end{aligned}$$

The primary and control routines listed in Tables 1 and 2 provide all the facilities furnished by the previous starting routines while the store space occupied is reduced from 32 to 29 commands because the “+K” function is used. Since these routines can be inserted without using equipment such as stepping switches, not otherwise required for testing purposes, considerable economy in equipment is possible, and the starting procedure is simplified.

References

- PEARCEY, T., and HILL, G. W. (1953*a*).—Programme design for the C.S.I.R.O. Mark I computer.
I. Computer conventions. *Aust. J. Phys.* **6** : 316–34.
PEARCEY, T., and HILL, G. W. (1953*b*).—Programme design for the C.S.I.R.O. Mark I computer.
II. Programme techniques. *Aust. J. Phys.* **6** : 335–56.

EIGEN OSCILLATIONS OF COMPRESSIBLE, IONIZED FLUIDS*

By R. E. LOUGHHEAD†

The frequencies of eigen oscillations of regions of uniform magnetic field may be derived very simply from the hydromagnetic equations for a compressible fluid of infinite electrical conductivity in the cases where the regions are bounded by cylindrical and plane-parallel surfaces.

For small oscillations these equations may be written in the forms :

$$\frac{\partial \mathbf{h}}{\partial t} = \text{curl} (\mathbf{v} \times \mathbf{H}), \quad \dots\dots\dots (1)$$

$$4\pi\mu \frac{\partial \mathbf{v}}{\partial t} = (\text{curl } \mathbf{H}) \times \mathbf{h} + (\text{curl } \mathbf{h}) \times \mathbf{H} - 4\pi \text{grad } p, \quad \dots\dots\dots (2)$$

$$\frac{\partial \rho}{\partial t} = -\text{div} (\mu \mathbf{v}), \quad \dots\dots\dots (3)$$

$$\frac{\partial p}{\partial t} = q^2 \frac{\partial \rho}{\partial t}, \quad \dots\dots\dots (4)$$

where \mathbf{h} is the perturbation in the steady magnetic field \mathbf{H} , \mathbf{v} the corresponding change in the fluid velocity, ρ the variation in the steady mass density μ , p the variation in the gas pressure, and q the velocity of sound in the fluid.

* Manuscript received June 1, 1955.

† Division of Physics, C.S.I.R.O., University Grounds, Sydney.

Suppose a uniform magnetic field H_z is confined within a circular cylinder of radius R whose central line coincides with the z -axis. In cylindrical co-ordinates the equations for radial oscillations may be written in the forms :

$$i\omega h_z = -\frac{1}{r} \cdot \frac{\partial}{\partial r}(rv_r H_z), \quad \dots\dots\dots (5)$$

$$4\pi\mu i\omega v_r = -\frac{\partial}{\partial r}(H_z h_z + 4\pi q^2 \rho), \quad \dots\dots\dots (6)$$

$$i\omega \rho = -\frac{1}{r} \frac{\partial}{\partial r}(r\mu v_r), \quad \dots\dots\dots (7)$$

in which the variables are taken to be proportional to $e^{i\omega t}$. The boundary conditions across the surface $r=R$ follow from equations (5) . . . (7), namely,

$$\left. \begin{aligned} \Delta(v_r H_z) &= 0, \\ \Delta(H_z h_z + 4\pi q^2 \rho) &= 0, \\ \Delta(\mu v_r) &= 0. \end{aligned} \right\} \quad \dots\dots\dots (8)$$

Inside the cylinder, where the magnetic field has the constant value

$$H_z = H_1, \quad \dots\dots\dots (9)$$

the variation of v_r is found to be governed by the equation

$$\frac{\partial^2 v_r}{\partial r^2} + \frac{1}{r} \frac{\partial v_r}{\partial r} + \left(\frac{\omega^2}{V^2} - \frac{1}{r^2} \right) v_r = 0, \quad \dots\dots\dots (10)$$

where

$$V = \left(\frac{H_1^2}{4\pi\mu_1} + q^2 \right)^{\frac{1}{2}} \quad \dots\dots\dots (11)$$

is the velocity of hydromagnetic disturbances propagated at right angles to the field H_z , μ_1 being the fluid density inside the cylinder. The solution of (10), finite on the axis $r=0$, is

$$v_r = AJ_1\left(\frac{\omega r}{V}\right), \quad \dots\dots\dots (12)$$

where $J_1(x)$ is Bessel's function of the first kind of order unity and A is an arbitrary constant. Outside the magnetic region v_r obeys the equation

$$\frac{\partial^2 v_r}{\partial r^2} + \frac{1}{r} \frac{\partial v_r}{\partial r} + \left(\frac{\omega^2}{q^2} - \frac{1}{r^2} \right) v_r = 0, \quad \dots\dots\dots (13)$$

whose solution, finite at infinity, takes the form

$$v_r = BJ_1\left(\frac{\omega r}{q}\right) + CY_1\left(\frac{\omega r}{q}\right). \quad \dots\dots\dots (14)$$

In this equation $Y_1(y)$ is Bessel's function of the second kind of order unity and B and C are arbitrary constants.

The boundary conditions (8) impose three linear relations on the constants A, B , and C and the condition for the existence of non-trivial solutions reduces to the equation

$$J_1\left(\frac{\omega R}{V}\right) \cdot \frac{\omega R}{q} \left\{ J_1\left(\frac{\omega R}{q}\right) Y_0\left(\frac{\omega R}{q}\right) - J_0\left(\frac{\omega R}{q}\right) Y_1\left(\frac{\omega R}{q}\right) \right\} = 0. \tag{15}$$

However, by a standard result in the theory of Bessel functions (cf. Watson 1944)*

$$y\{J_1(y)Y_0(y) - J_0(y)Y_1(y)\} = 2/\pi,$$

and hence equation (15) is satisfied if

$$J_1\left(\frac{\omega R}{V}\right) = 0. \tag{16}$$

Equation (16) has an infinity of real roots corresponding to the zeros of the Bessel function of the first kind of order unity. Let j_n ($n=1, 2, 3, \dots$) denote the infinity of real roots of the equation

$$J_1(x) = 0,$$

then the eigen-frequencies of the various modes of radial oscillation of the magnetic cylinder are given by

$$\omega_n = \frac{j_n}{R} \left\{ \frac{H_1^2}{4\pi\mu_1} + q^2 \right\}^{\frac{1}{2}}, \tag{17}$$

where $n=1, 2, 3, \dots$

In the second case consider a uniform magnetic field H_z confined within a pair of parallel planes $x=\pm a$, referred to Cartesian axes Ox, y, z . Then, by a similar analysis, the eigen-frequencies of unidimensional oscillations in the x -direction at right angles to the magnetic field are found to be given by

$$\omega_n = \frac{n\pi}{2a} \left\{ \frac{H_1^2}{4\pi\mu_1} + q^2 \right\}^{\frac{1}{2}}, \tag{18}$$

where $n=1, 2, 3, \dots$

The analysis may be generalized to cover the situation in which the magnetic field takes the values $H_z=H_1$ for $r<R$ or $-a<x<a$, and $H_z=H_2$ for $r>R$ or $|x|>a$, respectively. It is found that the eigen-frequencies given by (17) and (18) remain unchanged provided, of course, that $H_1 \neq H_2$.

* WATSON, G. N. (1944).—"Theory of Bessel Functions." p. 77. (Cambridge Univ. Press.)

A PROPORTIONAL COUNTER WITH GRID CONTROL*

By R. J. NORMAN†

Introduction

Where accurate energy loss measurements are to be attempted, counter proportionality must be preserved over the entire active length of the anode wire. This may be endangered by attachment to electronegative impurities in the counter gas ; or by the counter end effect, which causes a serious loss of gas amplification at the wire terminations (Curran, Angus, and Cockroft 1949 ; Rossi and Staub 1949).

The end effect may be considerably reduced either by using large length/diameter ratios (Curran, Cockroft, and Insch 1950), when the contribution of a given end effect may be made small ; or by some improvement in the design of the wire terminating system. Two common methods are to use either an earthed guard tube (Curran, Angus, and Cockroft 1949), insulated from the centre wire, or an effective thickening of the centre wire itself (Coon and Nobles 1947). If the latter system is used, the terminating rod diameter must be chosen so that the gas amplification at its surface will be negligible compared with that along the wire. A recent technique (Curran and Cockroft 1951), ideally suited to short counters, employs field adjusting tubes located at each end of the counter, and these limit electron collection to a region free from end effects.

An alternative method of eliminating end effect is advanced, in which a grid system of fine wires completely surrounds the anode wire, as in Figure 1. Since the effective counter diameter is now reduced to that of the grid system, a large length/diameter ratio will be obtained, while the full counter diameter is still available for electron collection.

The Gridded Counter

Construction.—The proportional counter, shown sectionally in Figure 1, was designed for a cosmic ray shower experiment, in which a large collection volume was required together with rapid electron collection. In order to test the field configuration, a double scale model of the grid and anode structure was submerged in an electrolytic tank and a three-dimensional field plot taken, particular attention being given to the region near the wire termination. It was found that the lines of force became radial within 3 mm from the end of each guard

* Manuscript received June 7, 1955.

† Physics Department, University of Melbourne.

tube, indicating constancy of gas amplification over all but 6 mm of the 56.8 cm effective wire length.

The counter wall was rolled from 0.006 in. tinplate to minimize transition effects, and the joins within α -particle range of the collection volume soldered with pure tin instead of ordinary lead solder, which is badly contaminated (Bearden 1933). Accurately machined "Micallex" insulators carried the 24 stainless steel grid wires, and located the anode wire on the axis of the grid system to within 0.001 in. The central wire was made of spring steel, examined microscopically for defects, and reduced to a constant diameter of 0.0225 ± 0.0001 in. over its entire length. This was the greatest diameter for an operating potential below 3 kV, and permitted a more accurate wire treatment. Removable end pieces carried the anode and grid assemblies, and were screwed tightly into place prior to outgassing, after which the O-ring seals were inserted to make the counter vacuum tight. A filling tap is provided which is vacuum

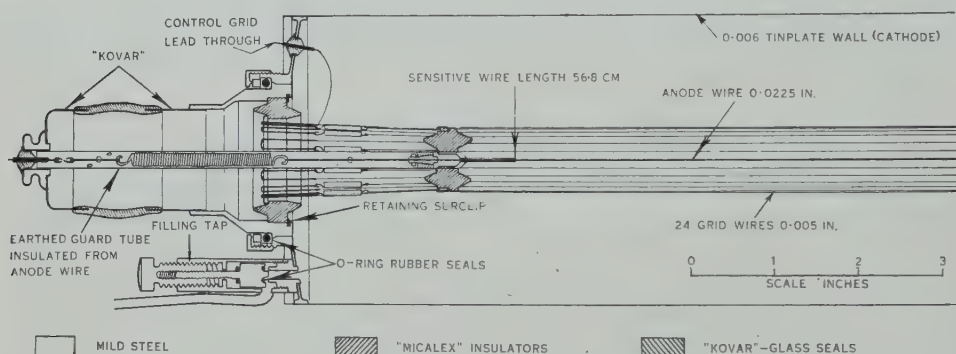


Fig. 1.—Sectional diagram of gridded proportional counter. Gas filling consists of 5 cm methane+argon to 81 cm Hg. Gas amplification is 400 for anode-grid voltage of 3 kV and anode-cathode voltage of 5 kV.

tight when closed, or when opened to the filling system. Gravitational sag is made the same in each wire by adjusting the tension with springs. The counter is enclosed in a thin metal shield to reduce radio-frequency interference, and to provide a safety precaution because of the high voltage present on the cathode.

For a gas amplification of 400, the grid is maintained at -3 kV with respect to the centre wire, and the cathode or case at -5 kV, allowing 2 kV for electron collection. The anode and guard rings are, of course, at earth potential.

Gas Filling.—The counter was outgassed for several days, at 160°C during the day; and left under static vacuum at room temperature overnight. When the vacuum remained better than 0.001 mm Hg for 24 hr, the counter was filled. Of the counter fillings tried, the most successful proved to be a mixture of Arc-argon and 6 per cent. commercial methane, dried with phosphorus pentoxide. Drift velocity data for these gases are available (English and Hanna 1953); the most important impurity being 0.2 per cent. oxygen in the Arc-argon.

The question of electron attachment to the oxygen impurity due to its known electronegative nature (Wilkinson 1950) may now be investigated by means of the control grid, since the grid conveniently partitions the counter into two sections; one being a high field region inside the grid wires which is most unfavourable for electron attachment, and the other being the large volume outside the grid system where attachment may occur. The latter occurs on forward collection where electrons liberated anywhere within the counter travel to the centre wire; whereas the former applies on reverse collection where a reverse field prevents electrons released outside the grid system from making any contribution.

A molybdenum filter on an X-ray tube concentrates the transmission radiation at its K absorption edge, providing a sufficiently homogeneous calibrating source to check the pulse distribution widths for both forward and reverse collection. These are shown in Figure 2 (a), the one for reverse collection being

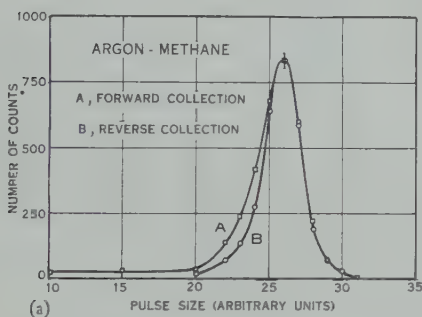


Fig. 2 (a).—Experimental pulse distributions obtained with transmission X-radiation through 0.017 in. of molybdenum, for a proportional counter filling mixture of unpurified commercial methane and Arc-argon. Percentage width at half height for *A* is 13 per cent. and for *B* 12 per cent.

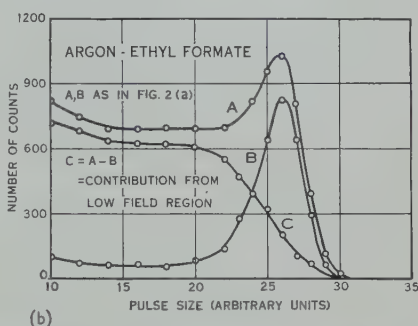


Fig. 2 (b).—Pulse distributions for a filling mixture of 6 per cent. ethyl formate with Arc-argon. Distribution width for the high field region inside the grid wires, *B* is 14 per cent. *A* represents the distribution from the entire volume and *C*, the contribution from the low field region outside the grid, has no peak, indicating a very high attachment rate.

12 per cent. width at half height, and for forward collection 13 per cent. This discrepancy in widths could be accounted for by the increased attachment outside the grid system, although this is by no means certain since some increase in width is to be expected from the X-rays which eject photoelectrons from the counter wall.

Two other counter fillings had previously been tried, one being Arc-argon with 6 per cent. ethylene. After working well for some time, it quickly deteriorated and produced a white deposit, presumably of some ethylene polymer, over the anode wire. The second filling contained 6 per cent. ethyl formate as quenching agent, and gave evidence of overwhelming attachment despite intense efforts to purify the liquid by distillation and outgassing at low temperature. The only conclusion possible was that ethyl formate is itself

slightly electronegative. This effect has not been observed with Geiger counters, possibly since they operate at a higher value of X/p than that outside the control grid system but comparable with that inside it, where X is the field strength in volt/cm and p the pressure in mm of Hg. The experimental pulse distributions for argon with ethyl formate are shown in Figure 2 (*b*), in which curve *B* is the contribution from inside the grid wires. Here the low distribution width provides evidence for only a small amount of attachment. In contrast to this, the distribution *C* obtained from the lower field region outside the grid wires exhibits no peak at all, and is indicative of an enormous attachment rate. The possibility of the collection field straddling an attachment peak may be discounted, since the field strength increases continuously, changing by an order of magnitude between cathode and grid.

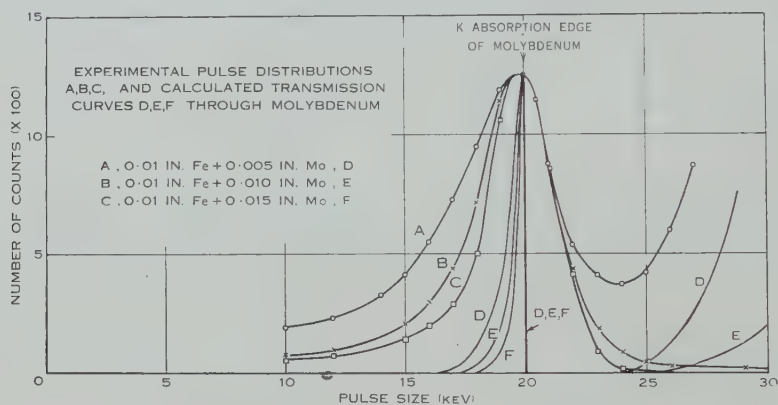


Fig. 3.—Theoretical and experimental curves for transmission X-radiation passing through the counter wall and shield (0.010 in. iron) plus various thicknesses of molybdenum. The energy scale is located by using the 46.7 keV γ -ray line from Ra D+E+F as a pulse size reference.

Calibration with Transmission X-radiation

A method of X-ray calibration with transmission radiation, utilizing the abrupt transmission discontinuity at the K absorption edge, has been found to yield excellent results. White radiation from an X-ray tube falls on a thin sheet of molybdenum, and the emergent radiation is grouped strongly below the K absorption edge at 20.0 keV. The fraction transmitted through various thicknesses of molybdenum, plus the 0.010 in. iron counter wall and shield is shown as a function of energy in Figure 3. As the thickness of molybdenum increases, the transmitted energy band becomes narrower and approaches K_{∞} , and the maximum thickness of molybdenum usable will then be governed by intensity considerations.

The experimental pulse distributions are also shown in Figure 3, and are accurately located on the energy scale by using the 46.7 keV γ -ray line from Ra D+E+F as an energy reference. The experimental distributions are seen to coincide on the high energy edge; and to exhibit a similar behaviour to the computed curves on the low energy side, which will be emphasized since the

counter detection efficiency at 15 KeV is just twice that at 20 keV. There will also be some contribution from fluorescent radiation at the smaller absorber thicknesses, since some higher energy radiation which penetrates the absorber will be accompanied by fluorescent radiation being continuously produced. The experimental peak occurs at 19.9 keV, which is equal to the energy of the absorption edge at 20.0 keV within the experimental error.

While there are objections to this type of calibration in principle, since the calibrating wavelengths are not homogeneous, in practice the width of the experimental pulse distributions is large compared with that of the transmitted energies. One may even assume that the latter are centred at the K absorption edge at 20 keV, the error so incurred being quite unimportant. It must also be emphasized that the calibration distribution width of 12 per cent. obtained with 0.017 in. molybdenum is not the smallest width available using the transmission method, but, when a tungsten anode X-ray tube is operated at 16 kV (r.m.s.) and 0.5 mA, a greater thickness that this will reduce the count rate through the single analyser channel to a value below 30 counts/sec. The homogeneity of the transmitted beam may be further improved by removing the contribution made by transmission fluorescent radiation. This fluorescence is produced by some of the higher energy X-rays which penetrate the absorber, and may be eliminated by a reduction in operating voltage.

The transmission method has several advantages. Firstly, a thin window, in the usual sense, is not necessary, provided that the counter wall itself does not cause too much attenuation. Of course the absorption edge of the wall material (6 keV) must be well below that of the transmitting element (20 keV). As a possible alternative, either some part or all of the counter wall may be constructed from the transmitting element. This can facilitate calibration of high or low pressure counters.

Secondly, since the emerging wavelengths are grouped strongly above the absorption edge, only one material is required for calibration. This contrasts favourably with either fluorescent X-radiation or K-capture calibration techniques, where the radiator or K-capture source emits all the characteristic wavelengths of the element concerned, and to produce a homogeneous beam for calibration we must use a selective filter to remove the $K\beta$ component. Consequently, if a selective filter is used, the number of calibrating energies available will be greatly reduced.

Finally, scattered radiation is completely eliminated since the counter may now be completely shielded with lead except for the transmission window, through which all the detected radiation must pass. The trouble experienced with scattering becomes very serious when fluorescent X-radiation is used for calibration. Here the contribution from scattered radiation makes the distribution peak shift with a change in X-ray tube voltage; even in the position of minimum scattering.

A debt of thanks is due to Mr. P. J. Walsh for the precision construction of the counter. The author also wishes to thank Professor L. H. Martin and Associate Professor C. B. O. Mohr for their continued interest in the project.

References

- BEARDEN, J. A. (1933).—*Rev. Sci. Instrum.* **4**: 271.
COON, J. H., and NOBLES, R. A. (1947).—*Rev. Sci. Instrum.* **18**: 44.
CURRAN, S. C., ANGUS, J., and COCKROFT, A. L. (1949).—*Phil. Mag.* **40**: 36.
CURRAN, S. C., and COCKROFT, A. L. (1951).—*Rev. Sci. Instrum.* **22**: 37.
CURRAN, S. C., COCKROFT, A. L., and INSCH, G. M. (1950).—*Phil. Mag.* **41**: 517.
ENGLISH, W. N., and HANNA, G. C. (1953).—*Canad. J. Phys.* **31**: 768.
ROSSI, B., and STAUB, H. H. (1949).—"Ionization Chambers and Counters." p. 92. (McGraw-Hill: New York.)
WILKINSON, D. H. (1950).—"Ionization Chambers and Counters." p. 41. (Cambridge Univ. Press.)

THE SOLUTION OF LINEAR SIMULTANEOUS EQUATIONS BY MATRIX ITERATION

By J. GUEST†

[Manuscript received April 22, 1955]

Summary

In a recent paper Stiefel presented a method designed for a high speed computer, for solving simultaneous linear algebraic equations of the type

$$\sum_{k=1}^n a_{ik}u_k + l_i = 0, \quad i=1, 2, 3, \dots, n.$$

The method proposed here arose from that paper. Moreover, since Stiefel fully examined the symmetric case, $a_{ik}=a_{ki}$, it seemed natural to develop the present theory for non-symmetric matrices also. Actually, Stiefel and Hestenes also touched on the non-symmetric problem but did not pursue the subject very far. A comparison between their method and that proposed here is given. As in Stiefel's theory the iteration ends at step n , which actually represents the exact solution provided no rounding-off errors have been committed. However, a different type of orthogonality and conjugate relation is used here as both D (i.e. the matrix $[a_{ik}]$) and its transpose D^* are operated with simultaneously. Formulae have been found for the characteristic polynomial of D and for its inverse.

I. INTRODUCTION

The problem is to solve a set of linear non-singular simultaneous algebraic equations

$$\sum_{k=1}^n a_{ik}u_k + l_i = 0, \quad (i=1, 2, 3, \dots, n). \quad \dots\dots\dots (1)$$

For values of n up to 10 this is probably best done by well-known methods such as Crout's. For n greater than 10, and especially when automatic equipment is available, iteration methods with accelerated convergence are superior. These methods have the advantage that inevitable rounding-off errors are kept in check and at the same time iteration methods are more suitable for digital computers.

The method outlined here is based essentially on the method of "minimal iterations" as described by C. Lanczos (1950). The problem dealt with here has been recently discussed by Hestenes and Stiefel (1952). When the matrix $[a_{ik}]$ is non-symmetric it appeared advantageous to depart from his suggested procedure and an alternative method is investigated.

† Aeronautical Research Laboratories, Department of Supply, Melbourne.

II. THE DEFINITIONS OF THE FUNDAMENTAL VECTORS AND PARAMETERS

The following six vectors and two parameters are introduced to start with :

$$\left. \begin{aligned} p_k &= -r_k + \varepsilon_{k-1} p_{k-1}, & k \geq 1, \quad p_0 &= -r_0, \\ v_{k+1} &= v_k + \lambda_k p_k, \\ r_{k+1} &= r_k + \lambda_k D p_k, \end{aligned} \right\} \dots\dots (2)$$

$$\left. \begin{aligned} p_k^* &= -r_k^* + \varepsilon_{k-1} p_{k-1}^*, & k \geq 1, \quad p_0^* &= -r_0^*, \\ v_{k+1}^* &= v_k^* + \lambda_k p_k^*, \\ r_{k+1}^* &= r_k^* + \lambda_k D^* p_k^*, \end{aligned} \right\} \dots\dots\dots (3)$$

$$\varepsilon_{k-1} = \frac{(r_k, D^* p_{k-1}^*)}{(p_{k-1}, D^* p_{k-1}^*)} = \frac{(r_k^*, D p_{k-1})}{(p_{k-1}^*, D p_{k-1})}, \quad \dots\dots\dots (4)$$

$$\lambda_k = - \frac{(r_k^*, p_k)}{(p_k, D^* p_k^*)} = - \frac{(r_k, p_k^*)}{(p_k^*, D p_k)}, \quad \dots\dots\dots (5)$$

where p_k is called the k th direction vector, v_k is called the k th solution vector, and r_k is called the k th residue vector. The above three vectors operate on D alone whilst p^* , v^* , and r^* operate only on D^* and carry the same names. Finally, ε_k and λ_k are, as will be shown shortly, suitable orthogonality parameters.

With the above definitions it is now possible to develop an algorithm to solve a system of n equations in n unknowns.

III. THE SOLUTION OF LINEAR EQUATIONS

It is required to solve (1) or, in matrix notation,

$$Du + l = 0, \quad \dots\dots\dots (6)$$

where D is the square matrix with elements a_{ik} ,

$$\begin{aligned} u &\equiv (u_1, u_2, u_3, \dots, u_n), \\ l &\equiv (l_1, l_2, l_3, \dots, l_n). \end{aligned}$$

In order to solve (6) it appears desirable to treat simultaneously

$$D^* u^* + l^* = 0, \quad \dots\dots\dots (7)$$

where D^* is the transposed matrix of D ,

u^* is a different solution vector (i.e. the one associated with D^*) and usually of no interest,

l^* is a conveniently chosen vector.

The first step in the analysis is to make a guess for u . This first approximation to the solution vector u is denoted by v_0 ; whilst successive approximations will be denoted by v_k . It then follows that

$$Dv + l = r, \quad \dots\dots\dots (8)$$

where r is called the residue vector. Likewise for (7) it follows that

$$D^* v^* + l^* = r^*. \quad \dots\dots\dots (9)$$

Using the definitions of v_{k+1} and v_{k+1}^* of (2) and (3) it follows, using (8), that

$$r_k = Dv_k + l,$$

and also

$$r_{k+1} = Dv_{k+1} + l, \quad \dots \dots \dots (10)$$

whence on subtraction

$$\begin{aligned} r_{k+1} - r_k &= D(v_{k+1} - v_k) \\ &= \lambda_k Dp_k \text{ by equation (2), } \dots \dots (11) \end{aligned}$$

which shows that, once v_k and p_k are defined, relation (11) is a direct consequence, that is, of the six defining vectors only four are independent.

Likewise, therefore,

$$r_{k+1}^* = r_k^* + \lambda_k D^* p_k^*. \quad \dots \dots \dots (12)$$

The λ_k are to be chosen in such a manner that successive residuals r_{k+1} will be orthogonal to p_j^* for $j=0, 1, 2, \dots, k$. It will be shown that this can actually be achieved by orthogonalizing r_{k+1} merely against p_k^* .

To fix λ_k it follows therefore that

$$\left. \begin{aligned} (r_{k+1}, p_k^*) &= 0, \\ (r_{k+1}^*, p_k) &= 0. \end{aligned} \right\} \quad \dots \dots \dots (13)$$

Using (11) and (12) this gives the first orthogonality parameter

$$\lambda_k = -\frac{(r_k^*, p_k)}{(p_k, D^* p_k^*)} = -\frac{(r_k, p_k^*)}{(p_k^*, Dp_k)}, \quad \dots \dots \dots (14)$$

provided $(p_k, D^* p_k^*)$ is non-vanishing. If the denominator vanishes then either r_k is orthogonal to p_k^* or it is required to start with a new vector v_0^* . It should be remembered, however, that it is only for very exceptional v_0^* that the above inner product would actually vanish. Of course it is still undesirable for this product to be very small. To be on the safe side in the choice of v_0^* one should try to choose a vector which is a linear combination of all the eigenvectors of D^* ; therefore it is usually best to choose for v_0^* a vector like

$$v_0^* = \{1, 1, 1, \dots, 1\}.$$

In order to fix ε_{k-1} , use the defining equations of p_k and p_k^* and postulate that

$$(r_k, p_{k-2}^*) = 0 = (r_k^*, p_{k-2}).$$

Post-multiplying (11) by p_k^* , it follows that

$$0 = (r_{k+1}, p_{k-1}^*) = (r_k, p_{k-1}^*) + \lambda_k \{-(Dr_k, p_{k-1}^*) + \varepsilon_{k-1}(Dp_{k-1}, p_{k-1}^*)\}.$$

Using (13) it follows that

$$\varepsilon_{k-1} = \frac{(r_k, D^* p_{k-1}^*)}{(p_{k-1}, D^* p_{k-1}^*)} = \frac{(r_k^*, Dp_{k-1})}{(p_{k-1}^*, Dp_{k-1})}, \quad \dots \dots \dots (15)$$

and therefore also

$$(p_k, D^* p_{k-1}^*) = 0, \quad \dots \quad (16)$$

provided

$$(p_{k-1}, D^* p_{k-1}^*) \neq 0.$$

The remarks made above on the vanishing of this product are still applicable here.

It is now possible to prove two fundamental theorems.

Theorem I

The system of residue vectors

$$\{r_0, r_1, r_2, \dots, r_{n-1}\}$$

is mutually orthogonal to

$$\{r_0^*, r_1^*, r_2^*, \dots, r_{n-1}^*\},$$

that is,

$$(r_i, r_j^*) = 0 = (r_i^*, r_j),$$

where $i, j = 0, 1, 2, \dots, n-1$ and $i \neq j$.

Theorem II

The system of direction vectors

$$\{p_0, p_1, p_2, \dots, p_{n-1}\}$$

is mutually conjugate to

$$\{p_0^*, p_1^*, p_2^*, \dots, p_{n-1}^*\},$$

that is,

$$(p_i, D^* p_j^*) = 0 = (p_j^*, D p_i),$$

where $i, j = 0, 1, 2, \dots, n-1$ and $i = j$.

These theorems will be proved by induction. Let it be assumed that Theorem II be true for $n=k$, that is,

$$(p_k, D^* p_{k-1}^*) = 0, \quad \dots \quad (17)$$

$$(p_k, D^* p_{k-2}^*) = 0, \quad \dots \quad (18)$$

$$(p_k, D^* p_{k-3}^*) = 0, \quad \dots \quad (19)$$

:

:

$$(p_k, D^* p_0^*) = 0.$$

It is required to prove it to be true for $n=k+1$. First, it is useful to establish the following

Lemma

Prove that

$$(i) \quad (r_k^*, r_{k+1}) = 0, \quad \dots \quad (20)$$

$$(ii) \quad (p_k^*, r_{k+1}) = 0, \quad \dots \quad (21)$$

$$(iii) \quad (r_k^*, r_{k+1}) = 0. \quad \dots \quad (22)$$

(ii) follows from the definition of ε_{k-1} .

Also, taking the defining equation of p_k^* and forming the scalar product with r_{k+1} gives

$$(r_{k+1}, r_k^*) = -(r_{k+1}, p_k^*) + \varepsilon_{k-1}(r_{k+1}, p_{k-1}^*) \\ = 0, \quad \dots\dots\dots (23)$$

using (13) and (11).

Finally, form the scalar product with r_{k-1}^* in (11):

$$(r_{k+1}, r_{k-1}^*) = (r_k, r_{k-1}^*) + \lambda_k(Dp_k, r_{k-1}^*) \quad \dots\dots\dots (23A) \\ = \lambda_k(Dp_k, r_{k-1}^*),$$

using (23).

But, by definition,

$$r_{k-1}^* = -p_{k-1}^* + \varepsilon_{k-2}p_{k-2}^*,$$

which, upon substitution in (23A), gives

$$(r_{k+1}, r_{k-1}^*) = -\lambda_k(Dp_k, p_{k-1}^*) + \lambda_k\varepsilon_{k-2}(Dp_k, p_{k-2}^*) \\ = 0, \quad \dots\dots\dots (24)$$

using (17) and (18), which proves (22).

Now, form the scalar product with $D^*p_{k-1}^*$ in

$$p_{k+1} = -r_{k+1} + \varepsilon_k p_k,$$

that is,

$$(p_{k+1}, D^*p_{k-1}^*) = -(r_{k+1}, D^*p_{k-1}^*) + \varepsilon_k(p_k, D^*p_{k-1}^*) \\ = -(r_{k+1}, D^*p_{k-1}^*).$$

But

$$\lambda_k D^*p_{k-1}^* = r_k^* - r_{k-1}^*,$$

whence

$$\lambda_{k-1}(p_{k+1}, D^*p_{k-1}^*) = -(r_{k+1}, r_k^*) + (r_{k+1}, r_{k-1}^*) \\ = 0, \quad \dots\dots\dots (25)$$

using (23) and (24).

Now

$$(p_{k+1}, D^*p_k^*) = 0 \text{ by (16),}$$

and

$$(p_{k+1}, D^*p_k^*) = 0 \text{ by (25).}$$

Likewise, it can be proved that p_{k+1} and p_{k-2}^* are mutually conjugate, and it can be shown at the same time that r_{k+1} and r_{k-2}^* are mutually orthogonal and so on until it is shown that p_{k+1} and p_0^* are mutually conjugate and r_{k+1} and r_0^* are mutually orthogonal. So, if the theorems be true for $n=k$, they will also be true for $n=k+1$. But the theorems are true for $k=1$, for

$$(r_1, r_0^*) = 0 \text{ by the choice of } p_0 \text{ and (13)}$$

and also

$$(p_1, D^*p_0^*) = 0 \text{ by (16).}$$

Hence the theorems are true for all n .

Theorem III

Asterisks can always be interchanged from one side of an inner product to the other.

Using the definition of p_k and p_k^* it follows immediately by using (13) that

$$(p_k, r_k^*) = (p_k^*, r_k).$$

It follows by induction that the stars are interchangeable in the product (p_k, r_j^*) for $j > k$ and it will be proved presently that this product vanishes for $j < k$ in either case. The orthogonality relations of Theorem I ensure that $(r_k, r_j^*) = (r_k^*, r_j)$. That $(p_k, p_j^*) = (p_k^*, p_j)$ can be shown by induction by using the definitions of p_k and p_j^* and the fact that $p_0 = -r_0$ and $p_0^* = -r_0^*$.

As regards the interchangeability of the stars in expressions like $(r_k, D^* p_j^*)$, the definition of p_k and induction again easily lead to the results :

$$\left. \begin{aligned} (r_k, D^* p_j^*) &= (r_k^*, D p_j) = (D r_k, p_j^*) = (D^* r_k^*, p_j), \\ (p_k, D^* p_j^*) &= (p_k^*, D p_j), \\ (r_k, D^* r_j^*) &= (r_k^*, D r_j). \end{aligned} \right\} \dots (26)$$

Therefore, it is always permissible to interchange asterisks from one side of an inner product to the other.

An interesting result of lesser importance is the following :

Theorem IV

The residue vector r_{k+1}^* is mutually conjugate to the system $\{r_i\}$ with $i = 0, 1, 2, \dots, k-1$.

This is easily proved with the help of Theorem II.

By definition

$$r_i = -p_i + \varepsilon_{i-1} p_{i-1}.$$

Forming the scalar product with r_{k+1}^* in the above and operating with D gives

$$\begin{aligned} (r_{k+1}^*, D r_i) &= -(r_{k+1}^*, D p_i) + \varepsilon_{i-1} (r_{k+1}^*, D p_{i-1}), \\ &= -\{-p_{k+1}^* + \varepsilon_k (p_k^*, D p_i)\} + \varepsilon_{i-1} \{-p_{k+1}^* + \varepsilon_k (p_k^*, D p_{i-1})\} \\ &= 0, \end{aligned}$$

since p_k^* is mutually conjugate to the system $\{p_i\}$ $i = 0, 1, 2, \dots, k-1$ by Theorem II. Hence the result.

Theorem V

For a system of n unknowns this iteration method will give the exact solution in n steps.

Every r_k is a linear combination of $r_0, D r_0, D^2 r_0, \dots, D^{k-1} r_0$ and similarly r_k^* is a linear combination of $r_0^*, D^* r_0^*, (D^*)^2 r_0^*, \dots, (D^*)^{k-1} r_0^*$.

If r_0 and r_0^* have components along all eigenvectors and principal vectors then these chains of vectors will be linearly independent up to $k = n$. Since r_n is orthogonal to all elements of the chain $r_0^*, D^* r_0^*, (D^*)^2 r_0^*, \dots, (D^*)^{n-1} r_0^*$, it must therefore be zero. Consequently the problem must be solved in n steps.

IV. THE INVERSE OF A SQUARE MATRIX D

We shall show that the general element of D , i.e. the inverse of D , is

$$a_{ij} = \sum_{k=0}^{n-1} \frac{p_{ki} p_{kj}^*}{(p_k, D^* p_k^*)}, \quad \dots \dots \dots (27)$$

where p_k is the direction vector defined in Section II and p_{ki} its i th component.

This is done by formally solving

$$Du_i = e_i$$

(where e_i is the unit vector with a 1 in the i th place and zeros elsewhere) for $i=1, 2, 3, \dots, n$.

The solution

$$Du = -f$$

can always be expressed as a linear combination of $p_0, p_1, p_2, \dots, p_{n-1}$ in the form

$$u = \alpha_0 p_0 + \alpha_1 p_1 + \alpha_2 p_2 + \dots + \alpha_n p_{n-1}. \quad \dots \dots \dots (28)$$

Assume this to be the case, for it will always happen unless the iteration procedure terminates before n steps, i.e. very exceptionally. Now find the α_i by using the biconjugate relation of the p_k and p_k^* (i.e. Theorem II).

For, by post-multiplying (28) by $D^* p_j^*$ it follows that

$$\alpha_j (p_j, D^* p_j^*) = (u, D^* p_j^*) = (Du, p_j^*) = -(f, p_j^*), \quad \dots \dots \dots (29)$$

whence

$$\alpha_j = -\frac{(f, p_j^*)}{(p_j, D^* p_j^*)}. \quad \dots \dots \dots (30)$$

Let now

$$-f = e_i,$$

with $i=1, 2, 3, \dots, n$ in turn, and let the corresponding u be u_i , i.e. $Du_i = e_i$. Then the matrix whose columns are u_i is really the inverse of D . If the j th component of p_k is p_{kj} then the j th component of u_{ij} is given by

$$\sum_{\rho=0}^{n-1} \alpha_{\rho} p_{\rho j} = \sum_{\rho=0}^{n-1} \frac{(e_i, p_{\rho}^*)}{(p_{\rho}, D^* p_{\rho}^*)} p_{\rho j}.$$

But $(e_i, p_{\rho}^*) \equiv p_{\rho i}^*$, which, on substitution, gives the right-hand side of expression (27).

V. THE CHARACTERISTIC EQUATION OF D

Let $q_k(x)$ be a polynomial of degree k , and let $q_{k+1}(x)$ be related to $q_k(x)$ in the same way as r_{k+1} is related to r_k . Thus

$$r_{k+1} = (1 + \gamma_k - \lambda_k D) r_k - \gamma_k r_{k-1} \quad \dots \dots \dots (31)$$

is replaced by

$$q_{k+1}(x) = (1 + \gamma_k - \lambda_k x) q_k(x) - \gamma_k q_{k-1}(x), \quad \dots \dots \dots (32)$$

where

$$\gamma_k = \frac{\lambda_k \varepsilon_{k-1}}{\lambda_{k-1}}, \quad \dots \dots \dots (33)$$

and the λ_k and ε_k are as defined earlier. Equation (31) follows from the recurrence relations

$$r_{k+1} = r_k + \lambda_k D p_k, \quad \dots \quad (11 \text{ bis})$$

$$p_k = -r_k + \varepsilon_{k-1} p_{k-1}, \quad \dots \quad (2 \text{ bis})$$

and the transformed version of the first by replacing $k+1$ by k , that is,

$$D p_{k-1} = \frac{1}{\lambda_{k-1}} \{r_k - r_{k-1}\}, \quad \dots \quad (11 \text{ ter})$$

on substituting p_k of (2 bis) into (11 bis) and by subsequently substituting $D p_{k-1}$ as given in (11 ter).

It is seen that $r_n = 0$, but also $r_n = \varphi_n(D) r_0$, where φ_n is a polynomial of degree n in D . If r_0 has components in the directions of all the eigenvectors of D then by the argument of Silberstein (1952) it follows that

$$\varphi_n(D) = 0, \quad \dots \quad (34)$$

and hence, by the Cayley-Hamilton theorem,

$$\varphi_n(x) = 0 \quad \dots \quad (35)$$

is the characteristic equation of D . By definition (32)

$$q_n(x) = \varphi_n(x).$$

Hence the characteristic equation of D is given by

$$q_n(x) = 0,$$

with

$$q_0(x) = 1,$$

$$q_1(x) = 1 - \lambda_0 x,$$

and the later $q_k(x)$ are developed by the recurrence relation as given by equation (32).

VI. AN ILLUSTRATIVE EXAMPLE

It is desired to solve the following system:†

$$\begin{bmatrix} 22 & -14 & 2 \\ -7 & 15 & -5 \\ 2 & -10 & 6 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 0. \quad \dots \quad (36)$$

The corresponding transpose equation is

$$\begin{bmatrix} 22 & -7 & 2 \\ -14 & 15 & -10 \\ 2 & -5 & 6 \end{bmatrix} \begin{bmatrix} v_1^* \\ v_2^* \\ v_3^* \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 0. \quad \dots \quad (37)$$

For convenience, start with $v_0 = 0$ and $v_0^* = 0$. This gives now rise to the following system of vectors which are recorded in Table 1.

† This relates to the deflection of a clamped square plate. The finite difference equivalent of the governing differential equation $\nabla^4 w - p/D = 0$ had to be satisfied at nine equally spaced inner points. This made the original matrix of order 9×9 . By symmetry considerations this was condensed to the above 3×3 non-symmetric matrix. For convenience the constant $625p/Da^4$ was put equal to one.

TABLE I
ALGORITHM FOR SOLVING THREE LINEAR SIMULTANEOUS EQUATIONS IN THREE UNKNOWNNS

k	r_k^*	p_k	$D^*p_k^*$	r_k	p_k^*	Dp_k	v_k	v_k^m
0	-1	+1	+17	-1	+1	+10	0	0
	-1	+1	-9	-1	+1	+3	0	0
	-1	+1	+3	-1	+1	-2	0	0
1	+3.636363636	+0.669421488	-63.074380158	+1.727272727	-1.239669421	-13.487603302	+0.2727272727	+0.2727272727
	-3.454545455	+2.578512397	+79.338842974	-0.1818181818	+5.851239670	+14.280991739	+0.2727272727	+0.2727272727
	-0.1818181818	+3.942148760	-16.264402810	-1.545454545	+2.578512397	-0.793388434	+0.2727272727	+0.2727272727
2	-0.9801762125	-0.4136505650	+10.017659958	+0.7400881057	+0.3756622489	-7.563896053	+0.3217235683	+0.1819933920
	+2.3524229095	+0.3939529208	-24.042383900	+0.8634361242	+0.5008830003	-8.824545403	+0.4614537446	+0.7009911896
	-1.372246697	+3.525878634	+14.024723948	-1.6035242294	+2.629635742	+16.388441466	+0.5612610133	+0.4614537446
3	0	0	0	0	0	0	+0.28125	+0.21875
	0	0	0	0	0	0	+0.5	+0.75
	0	0	0	0	0	0	+0.90625	+0.71875

$\epsilon_0 = +2.396694215$ $\lambda_0 = +0.2727272727$

$\epsilon_1 = +0.4876412642$ $\lambda_1 = +0.07319199709$

$\epsilon_2 = 0$ $\lambda_2 = +0.09784482755$

It is seen that by operating simultaneously with D and D^* not only the desired solution for the \mathbf{v}_k is obtained but also the one for \mathbf{v}_k^* as well. This is as it should be because of the mutual orthogonality relation with the residue vectors; \mathbf{r}_3^* had to be zero here, hence \mathbf{v}_3^* gave the exact solution to (37).

Furthermore, now that all ε_i and λ_i have been computed it is an easy matter to obtain the inverse of D .

Let the inverse matrix be given by

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

Then find successively (using the main result of Section IV):

$$\begin{aligned} a_{11} &= \sum_{\mu=0}^2 \frac{p_{\mu 1} p_{\mu 1}^*}{(p_{\mu}, D^* p_{\mu}^*)} = \frac{1}{11} - \frac{0.8298612384}{98.23591284} - \frac{0.1553929015}{35.83409646} \\ &= +0.078125, \\ a_{12} &= \sum_{\mu=0}^2 \frac{p_{\mu 1} D^* p_{\mu 2}^*}{(p_{\mu}, D^* p_{\mu}^*)} = \frac{1}{11} + \frac{3.916945567}{98.23591284} - \frac{0.2071905361}{35.83409646} \\ &= +0.125. \end{aligned}$$

Similarly, it is found that

$$\begin{aligned} a_{13} &= +0.078125, \\ a_{21} &= +0.0625, \\ a_{22} &= +0.25, \\ a_{23} &= +0.1875, \\ a_{31} &= +0.078125, \\ a_{32} &= +0.375, \\ a_{33} &= +0.453125. \end{aligned}$$

Finally, let the characteristic polynomial for the above matrix be computed (using the main result of Section V).

It is found that

$$\begin{aligned} q_0(x) &= 1, \\ q_1(x) &= 1 - 0.2727272727x, \\ q_2(x) &= 1 - 0.5213381059x + 0.01996145375x^2, \\ q_3(x) &= 1 - 0.78125x + 0.083984375x^2 - 0.001953125x^3, \end{aligned}$$

that is,

$$x^3 - 43x^2 + 400x - 512 = 0,$$

whilst

$$\begin{aligned} \gamma_1 &= +0.6432023988, \\ \gamma_2 &= +0.6518906069. \end{aligned}$$

VII. A COMPARISON WITH STIEFEL'S METHOD

Hestenes and Stiefel (1952) have also briefly discussed the non-symmetrical case. They arrived at the following iteration formulae :

$$\left. \begin{aligned} r_0 &= k - Ax_0, & p_0 &= A^* r_0, \\ a_i &= \frac{|A^* r_i|^2}{|Ap_i|^2}, \\ v_{i+1} &= v_i + a_i p_i, \\ r_{i+1} &= r_i - a_i A p_i, \\ b_i &= \frac{|A^* r_{i+1}|^2}{|A^* r_i|^2}, \\ p_{i+1} &= A^* r_{i+1} + b_i p_i. \end{aligned} \right\} \dots\dots\dots (38)$$

It was next attempted to solve the following system of six equations in six unknowns† by

- (i) Stiefel's method,
- (ii) the present method.

$$\begin{bmatrix} +22 & -16 & +2 & +2 & 0 & 0 \\ -8 & +23 & -7 & -8 & +3 & 0 \\ +1 & -7 & +13 & +2 & -6 & +1 \\ +2 & -16 & +4 & +20 & -14 & +2 \\ 0 & +3 & -6 & -7 & +14 & -5 \\ 0 & 0 & +2 & +2 & -10 & +6 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \end{bmatrix} = \begin{bmatrix} +1 \\ +1 \\ +1 \\ +1 \\ +1 \\ +1 \end{bmatrix} \dots (39)$$

To save space only the successive v_k vectors will be shown in Table 2. v_0 was taken as 0 in both cases.

The correct solution as given by Crout's method is (with a possible error of one in the last figure) :

$$\begin{bmatrix} +0.385284810 \\ +0.837816454 \\ +1.10007911 \\ +1.86431962 \\ +2.47587025 \\ +3.30498417 \end{bmatrix}.$$

The reason for the slower convergence in Stiefel's case is due to the fact that Stiefel's procedure is essentially a procedure with the matrix D^*D . The eigenvalues of this matrix are necessarily more widely spaced than for D alone. Consequently, the rate of convergence will be adversely affected if the constant vector l has large components along the largest and smallest eigenvectors.

† The problem with which this equation is associated is the same as that described in the footnote of Section V, the subdivision now having 25 inner points. By symmetry only six prove to be independent.

TABLE 2
SUCCESSIVE \mathbf{v}_k VECTORS

\mathbf{v}_k	Stiefel's Method	Present Method
\mathbf{v}_1	+0.009353718895	+0.4285714286
	-0.007152843861	+0.4285714286
	+0.004401750068	+0.4285714286
	+0.006052406344	+0.4285714286
	-0.007152843861	+0.4285714286
	+0.002200875034	+0.4285714286
\mathbf{v}_2	+0.06372821316	+0.4572531715
	+0.03593381053	+0.6580253718
	+0.009822136321	+0.6293436290
	+0.003034374902	+0.8014340864
	-0.03071552448	+0.7727523435
	+0.01537565739	+0.7440706006
\mathbf{v}_3	+0.06620838207	+0.3722517070
	+0.05309477145	+0.8530106900
	+0.01698714733	+0.8942678642
	+0.02239227478	+1.558693553
	-0.03424002185	+1.473973065
	+0.01012951991	+1.405684445
\mathbf{v}_4	+0.1161590155	+0.4946151738
	+0.1209167665	+0.8280575085
	+0.05414183470	+0.9873382239
	+0.1122802989	+1.709336742
	+0.00582644703	+2.062209289
	-0.04490672156	+2.115900105
\mathbf{v}_5	+0.1248919599	+0.3845229637
	+0.1409036315	+0.8383734999
	+0.0958567300	+1.122971474
	+0.1253287238	+1.841998113
	+0.0201016119	+2.466182473
	-0.05565182276	+3.294146561
\mathbf{v}_6	+0.3844214078	+0.3852848081
	+0.8380317059	+0.8378164566
	+1.098388256	+1.100079112
	+1.861448352	+1.864319617
	+2.473336988	+2.475870252
	+3.300624954	+3.304984174

Numerous checking facilities are available for either method. It is, however, pointless to carry out more than the most essential checks and these are :

- (i) column checks for all the included vectors,
- (ii) λ checks, e.g. $(r_k^*, r_0) = 0$,
- (iii) ε checks, e.g. $(p_k, D^*p_0^*) = 0$.

VIII. THE CORRECTION OF ROUNDING-OFF ERRORS

Rounding-off errors may become quite serious, in particular for large n . These types of errors can, however, be minimized by using an artifice due to Stiefel.

Let it be assumed that step i has just been completed in the computation and it is subsequently found that

$$(Dp_{i-1}, p_i^*) \neq 0, \quad \dots \quad (40)$$

but is fairly small of course. (If this is not so the error is due to the computer.)

It is now desirable to redefine λ_i and ε_i in such a manner that

$$(r_i, r_{i+1}^*) = 0, \quad \dots \quad (41)$$

that is, assuring that r_{i+1}^* will be orthogonal to the old r_i vector, and

$$(Dp_i, p_{i+1}^*) = 0, \quad \dots \quad (42)$$

that is, assuring that p_{i+1}^* will be orthogonal to the old Dp_i vector. It is necessary to prove the following

Lemma

$$(i) \quad (r_i, r_{i+1}^*) = (r_i, r_i^*) - \lambda_i(p_i, D^*p_i^*) + \varepsilon_{i-1}\lambda_i(p_{i-1}, D^*p_i^*), \dots \quad (43)$$

$$(ii) \quad \lambda_i(Dp_i, p_{i+1}^*) = -(r_{i+1}, r_{i+1}^*) + (r_i, r_{i+1}^*) + \varepsilon_i\lambda_i(Dp_i, p_i^*). \dots \quad (44)$$

Using the definition of r_i and post-multiplying with r_{i+1}^* gives

$$(r_i, r_{i+1}^*) = -(p_i, r_{i+1}^*) + \varepsilon_{i-1}(p_{i-1}, r_{i+1}^*). \dots \quad (45)$$

Now substitute for $r_{i+1}^* = r_i^* + \lambda_i D^*p_i^*$,

$$\begin{aligned} (r_i, r_{i+1}^*) &= -(p_i, r_i^*) - \lambda_i(p_i, D^*p_i^*) \\ &\quad + \varepsilon_{i-1}(p_{i-1}, r_i^*) + \varepsilon_{i-1}\lambda_i(p_{i-1}, D^*p_i^*) \\ &= (r_i, r_i^*) - \lambda_i(p_i, D^*p_i^*) + \varepsilon_{i-1}\lambda_i(p_{i-1}, D^*p_i^*), \dots \end{aligned} \quad (46)$$

using the definition of r_i^* . This proves (43).

Also by pre-multiplying the definition of p_{i+1}^* by Dp_i the following relation results:

$$(Dp_i, p_{i+1}^*) = -(Dp_i, r_{i+1}^*) + \varepsilon_i(Dp_i, p_i^*). \dots \quad (47)$$

Now substitute for $\lambda_i Dp_i = r_{i+1} - r_i$ in (47). This gives

$$\lambda_i(Dp_i, p_{i+1}^*) = -(r_{i+1}, r_{i+1}^*) + (r_i, r_{i+1}^*) + \varepsilon_i\lambda_i(Dp_i, p_i^*),$$

which proves (44).

Next introduce a λ'_i which is slightly different from λ_i and choose it such that

$$(r_i, r_{i+1}^*) = 0. \quad \dots \quad (41 \text{ bis})$$

Using result (i) of the lemma and also the fact that $(r_i, r_i^*)=(r_i^*, p_i)$ in (10) yields at once

$$\lambda_i'=\frac{+(r_i, r_i^*)}{(p_i, D^*p_i^*)-\varepsilon_{i-1}(p_{i-1}, D^*p_i^*)} \dots\dots\dots (48)$$

$$=\frac{\lambda_i}{d_i}, \dots\dots\dots (49)$$

where

$$d_i=1-\varepsilon_{i-1}\frac{(p_i^*, Dp_{i-1})}{(p_i, D^*p_i^*)} \dots\dots\dots (50)$$

There is now a refined λ_i' at our disposal which assures that $(r_i, r_{i+1}^*)=0$ or at any rate is much smaller than it had been before the correction term was applied.

Further, let an ε_i' be now introduced—slightly different from the old ε_i —which shall be chosen such that

$$(Dp_i, p_{i+1}^*)=0. \dots\dots\dots (42 \text{ bis})$$

Now equation (44) holds for *both* ε_i and ε_i' and also λ_i and λ_i' , i.e.†

$$\varepsilon_i\lambda_i(Dp_i, p_i^*)=+(r_{i+1}, r_{i+1}^*), \dots\dots\dots (51)$$

$$\varepsilon_i'\lambda_i'(Dp_i, p_i^*)= (r_{i+1}, r_{i+1}^*). \dots\dots\dots (52)$$

Hence using (51), (52), and (49) we obtain

$$\varepsilon_i'=\varepsilon_id_i. \dots\dots\dots (53)$$

Thus both λ_i and ε_i have been refined for rounding-off errors.

IX. CONCLUSIONS

The above method of solving systems of n equations in n unknowns seems to be well suited for an electronic high speed computer, since once a programme for an affine transformation has been devised the rest is quite straightforward. However, the method is not very fast. In fact, compared with one of the pivotal condensation methods the present approach requires nearly three times as many more multiplications. Against that should be weighed the undoubted advantage of having control of round-off errors. The method is therefore not suitable for desk machines for that reason. A good computer may complete a 10×10 matrix in about 8 working hours when using the usual Crout's method approach but would spend about five times that time on the above method. It is important to keep some checking facilities going when proceeding from one step to the next. It is considered desirable to carry all column checks, one bi-orthogonality test, and one test for the biconjugate relation. It will be found that the effect of rounding-off errors becomes rather appreciable as n increases, but this can

† The second term on the right-hand side must vanish by (41) for λ_i' or by the bi-orthogonality relations of the r_k and r_j^* for λ_i .

be overcome to a large extent by going beyond n steps. Lanczos (1950) suggests a test for orthogonality by adding to b_i a correction term ϵ_{ij} as defined by

$$\epsilon_{ij} = -\frac{(b_i, b_j)}{(b_j, b_j)} b_j,$$

if b_i is appreciably lacking in orthogonality with another vector b_j of whose orthogonality we are certain. This, however, has the obvious weakness that while $(b_i, b_j)=0$ now, the new b_i will disturb the previous orthogonality so that in fact nothing better has been gained in the end.

The present method, outlined above, is, in general, superior to Stiefel's as pointed out in Section VI, but some disadvantages of the method must also be mentioned.

- (i) As compared with Stiefel's method, its computing time is slightly longer.
- (ii) The method may fail altogether if

$$(p_k, D^* p_k^*)=0,$$

which is, however, rather unlikely.

X. ACKNOWLEDGMENT

The author would like to take this opportunity to thank the Chief Scientist, Department of Supply, for permitting that this paper be published in an outside Australian journal.

XI. REFERENCES

- HESTENES, M. R., and STIEFEL, E. (1952).—Method of conjugate gradients for solving linear systems. U.S. Nat. Bur. Stand. Rep. 1659.
- LANCZOS, C. (1950).—An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *J. Res. Nat. Bur. Stand., Wash.* **45**: 255–82.
- SILBERSTEIN, J. P. O. (1952).—On the method of minimal iterations. Dep. Supply Aust. Aeron. Res. Lab. Rep. SM 200.

THE STATISTICAL DESIGN OF AN EXPERIMENT TO TEST THE STIMULATION OF RAIN

By P. A. P. MORAN*

[*Manuscript received July 8, 1955*]

Summary

The problem of analysing statistically the results of a C.S.I.R.O. rainfall experiment over the catchment area of the Kiewa river is considered. It is shown that in order to obtain a significant result with any reasonable supposed increase in the rainfall within a few years it is necessary to use a control variable strongly correlated with the test variate. The powers of various tests using either streamflow or rainfall as test variates are calculated when the control variable is either streamflow in the Murray or rainfall in other areas. The general problem of designing such tests and the relationship between rainfall and streamflow are also considered.

I. INTRODUCTION

Since the latter part of 1954 rainfall stimulation using a silver iodide generator has been attempted by the C.S.I.R.O. Division of Radiophysics over the Kiewa area in Victoria. The generator is sited on Mount Stanley and the silver iodide is released whenever the wind is blowing into the 45° sector south to south-east of the generator. Decisive results in such an experiment cannot be expected for several years and the statistical analysis requires careful consideration. By making a preliminary analysis, however, we can determine how long such an experiment must be continued in order to have any specified probability of detecting a given increase in rainfall. Moreover, many experiments of this type have been carried out elsewhere without any idea of the statistical problems involved and have thus led to much wasted effort. It seems therefore worth while to discuss these problems at some length before the actual results of the experiment are obtained.

The statistical analysis involves taking some measured quantity X which it is hoped will be affected by the stimulation and deciding whether the observed difference between the mean values of X in the presence and absence of stimulation is larger than might be reasonably expected to have occurred by chance. Thus the analysis of the experiment is essentially a statistical one.

We must consider two characteristics of such a statistical test—its validity and its power. The test assumes a null hypothesis that there is no difference between the expected values of the observed variate in the presence and absence of stimulation. We then calculate a quantity, depending on the observed difference, whose statistical distribution is known under certain assumptions. The test will be a valid one if these assumptions are in fact true. In particular we have to make sure of the normality of the distribution of the test variate

* Australian National University, Canberra, A.C.T.

and of the absence of serial correlation and trend. If the latter are present a different form of analysis would be necessary.

We must also consider the power of the test, that is, the probability of obtaining a statistically significant result when a given increase in the expected rainfall occurs. Clearly, the greater the power of the test, the shorter the time necessary to make an adequate test of the method of stimulation. The power of the test will depend in part on the test criterion used and in part on the design of the experiment. The most effective way of increasing the power of the test is to use another variate (the "covariate") which is unaffected by the seeding and which is strongly correlated with the test variate. Regression analysis then enables most of the variability of the latter to be eliminated and the power of the test is then greatly strengthened. In fact, detailed analysis shows that, without the use of such a covariate, experiments to test rain stimulation would have to be impracticably long.

We also have to choose the test variate. In the present case there are two such variates which it would be natural to choose—the river flow of the Kiewa River, and the rainfall in the test area. We shall consider the analyses of these separately and then consider what can be said about their relative merits. In both cases we shall take annual values. It might, at first sight, be thought better to use monthly or even weekly values. It is easy to show, however, that so long as a sufficient number of observations are used to obtain a good estimate of the standard deviations and correlations involved, there is no increase in power by using shorter intervals. Moreover, shorter intervals produce many more troubles in the shape of seasonal variation, non-normality, and serial correlation.

It has, however, been pointed out to me by a meteorologist friend that only part of the annual rainfall is received from winds blowing in the directions considered. He estimates, roughly, that Bright and Omeo, for example, receive about 40 and 30 per cent. respectively of their annual rainfall from winds between north and north-west. Thus, assuming rainfall is not correlated with wind direction, to produce a 10 per cent. increase in annual rainfall, the seeding operation would have to increase rain by 25–33 per cent. on the occasions on which it is used. This reduces the power of the test relative to a given increase in rainfall. In what follows, in speaking of a given percentage increase we shall mean that percentage increase in the target area and not the maximum possible increase which might be obtained if the target area had been seeded from all directions. These considerations suggest that a more powerful test might be obtained by confining the data to cases when the wind was in the right direction. This might be done but would require a considerably more complicated analysis and would in any case not be applicable to the case where stream flow is the test variate. In the present paper we therefore confine ourselves to annual data.

II. ANALYSIS WITH RIVER FLOW AS TEST VARIATE

We define x_1 as the annual river flow at Kiewa. This is known from 1886 onwards but we shall only use the values from 1891 since we are going to correlate it with x_2 , the annual flow of the Murray at Jingellic, which is only known from

1891 onwards. Here and in what follows we consider the period of 58 years from 1891 to 1948. In the final analysis of the experiment we will have further years added to these and we give in this paper all the sums, sums of squares, and sums of products of the variates used in order to reduce the work of future analysis. For the 58 years used we have, in thousands of acre feet,

$$\begin{aligned}\Sigma x_1 &= 30125, & \bar{x}_1 &= 519.397, \\ \Sigma x_1^2 &= 19672017, & \text{S.D. } (x_1) &= 265.74. \\ \text{Coefficient of variation of } x_1 &= 0.51.\end{aligned}$$

Both the series x_1 and x_2 have somewhat skew distributions and they were therefore transformed by taking logarithms to the base 10 to four decimal places. We write $x_3 = \log_{10} x_1$ and $x_4 = \log_{10} x_2$.

Frequency distributions of x_3 and x_4 are shown in Tables 1 and 2.

TABLE 1
FREQUENCY DISTRIBUTION OF x_3

Values ..	2.2-	2.4-	2.6-	2.8-	3.0-	3.2-3.4	Total
Frequency ..	6	14	24	12	1	1	58

TABLE 2
FREQUENCY DISTRIBUTION OF x_4

Values ..	0.6-	0.8-	1.0-	1.2-	1.4-	1.6-1.8	Total
Frequency ..	1	4	19	26	7	1	58

A normal distribution was fitted to x_3 and the observed fit was good. This gives some confidence that residuals from the regression of x_3 on x_4 are normally distributed. Taking the values from 1891 to 1948 we get:

$$\begin{aligned}\Sigma x_3 &= 154.6147, & \Sigma x_4 &= 71.2300, \\ \Sigma x_3^2 &= 414.71706619, & \Sigma x_4^2 &= 89.46846034, \\ \Sigma x_3 x_4 &= 191.92945407, & r_{34} &= 0.9084.\end{aligned}$$

The high value of the correlation between x_3 and x_4 shows that x_4 is a useful and satisfactory covariate. Let z_1, \dots, z_n be the values of x_3 for the n years during which the area is seeded and w_1, \dots, w_n the corresponding values of x_4 . We then test whether the mean of z_1, \dots, z_n is significantly larger than that of x_3 when the effects of x_4 and w are removed (the above values extend only to 1948 and it would be as well to add the values for 1949-1953). The analyses of the experiment thus takes the form of a standard analysis of variance and covariance.

In the above data we already have 58 years of records so that r_{34} and the variances can be fairly well estimated. We can therefore use a rough calculation to determine, for any given number of years of seeding and any given supposed increase in streamflow, the probability of obtaining a significant result at a prescribed significance level, i.e. we can determine the power function of the test.

The standard deviation of x_3 is 0.2114 so that the standard deviation of x_3 corrected for x_4 is approximately $0.2114(1-r_{34}^2)^{\frac{1}{2}}=0.08838$. The standard deviation of the difference between the mean of the n years of seeding and the 63 years (1891–1953) which will be used in the final test is about

$$0.08838(n^{-1}+63^{-1})^{\frac{1}{2}}.$$

We are here considering a one-sided test so that we judge the observed difference of means to be significant at the 5 per cent. level if it exceeds 1.6449 times its standard deviation.

TABLE 3
PROBABILITIES OF SIGNIFICANT RESULT (5% POINT)

Increase (%)	Years		
	2	3	4
10	0.16	0.20	0.23
20	0.35	0.45	0.54
30	0.56	0.70	0.80
40	0.74	0.88	0.94
50	0.87	0.96	0.99
0	0.05	0.05	0.05

Increases of 10, 20, . . . , 50 per cent. in the mean streamflow will result in increases in the expected value of x_3 of 0.04139, 0.07918, 0.11394, 0.14613, and 0.17609 respectively. Using these values we calculate the probabilities given in Table 3 of judging the increase to be significant at the 5 per cent. point when the experiment lasts 2, 3, or 4 years. The method of calculation may be illustrated for the case of a 30 per cent. increase and a period of 3 years. The standard deviation used in the test is $0.08838(3^{-1}+63^{-1})^{\frac{1}{2}}=0.0522$ and multiplying this by 1.645 we get 0.0859. The mean of 3 years has a standard deviation (corrected for the regression) of 0.0510 and an expected value of 0.1139 so that its chance of exceeding 0.0859 is given by tables of the normal distribution and is 0.70 approximately. These results are only approximate as the above method of calculation is not that which would be used in the final analysis. Longer periods would be required to obtain the same probabilities of obtaining results significant at the 1 per cent. level.

These results are perhaps not very encouraging when it is realized that we are here working in the rather fortunate circumstances of having a covariate whose correlation with the test variate is about 0.9, but they emphasize the

fact that without a control variate experiments on data of this degree of variability would have to be very long indeed.

One possible way of improving the power is to use more than one control variate. An attempt to do this was made by taking a number of rainfall stations north and north-west of the target area. These were Albury, Rutherglen, Chiltern, Wangaratta, Benalla, and Tallangatta. All of these have uninterrupted records dating back to 1891 or earlier. As a control variate their sum, x_5 , was taken. For the 58 years 1891–1948 we have :

$$\begin{aligned}\Sigma x_5 &= 9065 \cdot 00, & \Sigma x_5^2 &= 1505538 \cdot 0460, \\ \Sigma x_3 x_5 &= 24548 \cdot 246409, & \Sigma x_4 x_5 &= 11499 \cdot 274353.\end{aligned}$$

Hence

$$\begin{aligned}r_{35} &= 0 \cdot 8053, \\ r_{45} &= 0 \cdot 8720.\end{aligned}$$

Hence

$$r_{35 \cdot 4} = 0 \cdot 0639,$$

which is not significant at the 10 per cent. point, and so the use of x_5 as a second covariate is not of any value.

We must also consider whether there is any trend in the series such as would arise from slow climate changes. That such climatic changes occur is well known (see for example Cornish (1954)). In the present problem such a trend, if it exists, may have little effect on the validity of the test because it would be likely to affect both the Kiewa and Jingellic flows more or less equally and we are only concerned with the residuals from the regression. Taking time as a new variate x_6 we get $r_{36} = 0 \cdot 0704$ and $r_{46} = 0 \cdot 0206$, neither of which is significant at the 10 per cent. point.

The above tests on correlation coefficients also assume the absence of serial correlation from year to year. The latter, if it exists, might seriously upset the test of significance of the experiment. The use of stream flow as a test variate, in contrast to rainfall, gives rise to a suspicion that such a correlation may exist because of the lagged run-off which results from single storms and also from the fact that snow may lie for some time. To ensure the validity of the final test we need to ensure that the residuals from the regression on the covariate are uncorrelated and we must therefore test these residuals for serial correlation. A one-sided test will be appropriate since such correlation as may exist is not likely to be negative. Notice that it is the serial correlation in the residuals rather than in the original series which is important.

An exact and asymptotically most powerful test for serial correlation in residuals from a regression is now available (Hannan 1955). However, a simpler approximate method which could be applied here is given by Durbin and Watson (1950, 1951). Such a test should be applied before the final analysis is made.

The tests for trend and serial correlation will answer, in part, any criticism that the records are not homogeneous. Such inhomogeneity might arise if (1) the earlier records were not obtained with the same standard of accuracy or

by the same method as the later records or if (2) there are local inhomogeneities in the series due to some specific meteorological factor which operates over a number of years and then disappears. The absence of trend in the series seems to indicate that (1) is not present to any significant effect and (2) also does not seem to appear and seems unlikely on other grounds. If there is a genuine effect of type (2) it should appear in the test of serial correlation. These statements apply also to the rainfall series.

It seems best therefore to go ahead and analyse the final experiment on the assumption that these disturbing factors are unimportant. If, however, convincing evidence can be produced to show that the series are sufficiently inhomogeneous to make the final analysis doubtful, another method of analysis can be used which, although slightly less powerful, will avoid these difficulties. This is done by truncating the series at 1945 (say) and testing whether the results for the test period (1955 onwards) are significantly larger than the means for 1946–53 (say) when the variances and covariances are estimated from 1891–1945 only. The above effects, if they exist, should then only have the effect of inflating the variances and the test will still be valid.

III. THE ANALYSIS WITH RAINFALL AS A TEST VARIATE

The only rainfall stations within the target area and not too far from Mount Stanley which have long and uninterrupted records are Bright (from 1881) and Omeo (from 1880). At first it was thought that the sum of these would make a good test variate after a logarithmic transformation and using only the years 1891–1948 since this was the period for which the other observations were available. This was done and the resulting variate was found to have an observed distribution which is well fitted by a normal distribution. As a control variate $x_9 = \log_{10} x_5$ was taken (x_5 being the sum of the rainfalls at the six stations previously mentioned). The correlation with x_9 was found to be 0.9525. This is remarkably high. Such a high correlation being somewhat unexpected it was decided to see what sort of correlation was to be expected between rainfall stations in such an area. As an experiment, therefore, the six control stations were split into two groups of three, Albury, Rutherglen, and Chiltern forming one group and Wangaratta, Benalla, and Tallangatta the other. The correlation between the sums of rainfall in these two groups, without using a transformation, was 0.9524 for the 58 years. This suggests that such high correlations are not unusual.

A correlation 0.9524 between the test and control variates is very satisfactory and suggests that a powerful test would result. However, both Dr. C. H. B. Priestley and Mr. E. B. Pender have pointed out to me that Omeo is in a rain shadow on the south-eastern side of the mountain range and is therefore, presumably, less sensitive to any increase in rainfall which might result from seeding from Mount Stanley. It was decided, therefore, to use the figures for Bright alone and define the annual rainfall there as x_7 . In any final analysis it would, of course, be interesting to see what happens to the Omeo rainfall as well.

To avoid skewness in the distribution we use a logarithmic transformation and define $x_8 = \log_{10} x_7$. For this variate we have the frequency distribution given in Table 4. As a control variate we take $x_9 = \log_{10} x_5$, which has a frequency distribution shown in Table 5.

TABLE 4
FREQUENCY DISTRIBUTION OF x_8 (1891-1948)

Values ..	1.35-	1.40-	1.45-	1.50-	1.55-	1.60-	1.65-	1.70-	1.75-	1.80-	1.85-	1.90
Frequency ..	2	3	3	8	10	9	12	9	1	0		1

TABLE 5
FREQUENCY DISTRIBUTION OF x_9 (1891-1948)

Values ..	1.90-	1.95-	2.00-	2.05-	2.10-	2.15-	2.20-	2.25-	2.30-	2.35-	2.40-	2.45
Frequency ..	2	3	1	6	9	12	12	4	5	2		2

These distributions are well fitted by normal distributions. We have :

$$\begin{aligned}\Sigma x_8 &= 93.2346, & \Sigma x_9 &= 126.4705, \\ \Sigma x_8^2 &= 150.485718, & \Sigma x_9^2 &= 276.45548583, \\ \Sigma x_8 x_9 &= 203.90216, \\ r_{89} &= 0.9307, & (1 - r_{89}^2) &= 0.1338.\end{aligned}$$

The standard deviation of x_8 is 0.10360 and so the standard deviation of x_8 corrected for x_9 is about 0.03790. The standard deviation of the mean of n years is then $0.03790n^{-\frac{1}{2}}$ and so the standard deviation of the difference of this mean and the mean of the 63 years (1891-1953) to be used in the final test will be $0.03790(n^{-1} + 63^{-1})^{\frac{1}{2}}$. As we are using a one-sided test, we take the observed

TABLE 6
PROBABILITIES OF OBTAINING SIGNIFICANT RESULTS FOR GIVEN PERCENTAGE
INCREASE IN RAINFALL

Increase (%)	Years		
	1	2	3
10	0.29	0.45	0.58
20	0.67	0.90	0.97
30	0.91	0.995	1.00

mean difference to be significant at the 5 per cent. level if it exceeds 1.645 times its standard deviation. For increases of 10, 20, and 30 per cent. in the expected values of x_7 we get the following approximate probabilities of obtaining a result significant at the 5 per cent. level (Table 6).

These results are approximate and the correct form of the final test is that of an analysis of variance and covariance together with a t -test.

IV. THE RELATIONSHIP BETWEEN RAINFALL AND RUN-OFF

The above calculations show that a given percentage increase in rainfall has a much higher probability of being detected than the same percentage increase in run-off. This is, of course, partly due to the fact that the correlation with the control variate is higher in the first case (0.9307) than in the second (0.9084). However, it is also due to the fact that the variation of the streamflow is larger compared with its mean than that of rainfall. This is clear from the coefficients of variation which are 0.51 for x_1 , the streamflow at Kiewa, and 0.24 for x_7 , the Bright rainfall. Thus a 20 per cent. increase in streamflow is smaller compared with the standard deviation of the latter than a similar increase in rainfall.

Since streamflow is the result of rainfall one would expect at first sight that the coefficients of variation of the two should be about the same. Furthermore, if one takes rainfall at only one or two stations in the target area one would expect the "error" resulting from taking this as a measure of the total precipitation on the area to have the effect of making the coefficient of variation of rainfall appear larger than that of streamflow. In the above results just the reverse is the case. This is due to the fact that run-off is an "excess" remaining after a certain amount of the rainfall has been lost by evaporation, transpiration, and seepage (see for example the discussion in Johnstone and Cross (1949, pp. 103-5)).

If this is the correct explanation it would be quite misleading to compare the powers of the two test variates by comparing the probabilities of obtaining a significant answer for the same annual percentage increase of each because a 20 per cent. increase (say) in rainfall will probably cause a much larger percentage increase in run-off. To make a useful comparison of the probabilities of detecting a 20 per cent. increase in rainfall when using rainfall or run-off would require a much more elaborate calculation. It may well turn out that run-off provides as powerful or nearly as powerful a test as rainfall and so both methods should be used in the final analysis.

The above-mentioned relationship between rainfall and run-off also suggests that the economic effects of a 20 per cent. (say) increase in rainfall may be much larger than those ascribable to a 20 per cent. increase in run-off.

V. CONCLUSIONS

(1) Both rainfall and streamflow will be useful test variates when associated with suitable covariates.

(2) The experiment should be continued for several years and will then provide a decisive test of the practicability of rain stimulation by these methods.

(3) Provided previous records exist for a long enough time in the past for the estimation of variances and correlation, annual values are to be preferred to monthly or weekly values.

VI. ACKNOWLEDGMENTS

The author is much indebted to Dr. J. L. Pawsey, F.R.S., Mr. E. B. Pender, and Dr. C. H. B. Priestley for a great deal of helpful criticism; and to Mr. L. J.

Dwyer and Mr. G. N. Alexander for help in studying data. Finally it should be stated that Mr. E. E. Adderley is also carrying out a statistical analysis of this experiment by different, and in some respects better, methods.

VII. REFERENCES

- CORNISH, E. A. (1954).—*Aust. J. Phys.* **7**: 334.
DURBIN, J., and WATSON, G. S. (1950).—*Biometrika* **37**: 409.
DURBIN, J., and WATSON, G. S. (1951).—*Biometrika* **38**: 159.
HANNAN, E. J. (1955).—*Biometrika* **42**: 133.
JOHNSTONE, D., and CROSS, W. P. (1949).—"Elements of Applied Hydrology." (Ronald Press: New York.)

PHOTODISINTEGRATION OF ^{16}O AND ^{63}Cu BY PHOTONS OF VARIABLE ENERGY

By J. G. CAMPBELL*

[Manuscript received August 31, 1955]

Summary

Proton capture reactions in light nuclei provide γ -radiation with energy controllable by variation of proton energy. The radiation from proton capture by ^7Li has been used to investigate the (γ, n) reactions in ^{16}O and ^{63}Cu over an energy range of 300 keV around 17.7 MeV. No resonant behaviour was observed in ^{16}O . A variation in the ^{63}Cu photoneutron cross section is interpreted as the combined effect of many unresolved levels.

I. INTRODUCTION

The possibility that nuclei might absorb γ -radiation through excitation of discrete energy levels, by a process inverse to monochromatic γ -emission, has been recognized for many years. Kuhn (1929) reported an unsuccessful attempt to detect selective scattering of ThC'' radiation by ^{208}Pb , which was based on the expectation that the radiation from the 2.62 MeV excited state of ^{208}Pb (resulting from β -emission by ThC'') would excite that same state in other identical ^{208}Pb nuclei. It is only recently, however, that the effect has been successfully detected.

Kuhn's experiment failed because the emission and absorption of the radiation must involve loss of energy to nuclear recoil, so that the energy absorbed is less than the excitation energy of the original level by an amount sufficient to take it off resonance. In recent experiments along the same lines this energy has been restored by causing the emitting nuclei to be moving towards the scattering nuclei so that the photon energy is increased by the Doppler effect. This motion has been achieved by direct mechanical means (Moon 1951; Davey and Moon 1953), by heating (Malmfors 1952), and by recoil from a previous emission (Ilakovac 1954). The resonant scattering so obtained has enabled the lifetimes of the states concerned to be determined.

Recent work on the activation curves resulting from disintegration of light nuclei by betatron radiation has disclosed abrupt changes in slope, or "breaks", which are interpreted as the effect of resonant absorption by discrete nuclear levels. The positions of several of the breaks for the reaction $^{16}\text{O}(\gamma, n)^{15}\text{O}$ found at Saskatchewan (Katz *et al.* 1954) have been confirmed by a group at Illinois (Penfold and Spicer, personal communication). Similar effects in the activation curve for $^{63}\text{Cu}(\gamma, n)^{62}\text{Cu}$ were reported by Phillips (1953), but neither of the groups at Saskatchewan and Illinois has been able to confirm this.

This paper reports the results of experiments carried out on these two reactions using the γ -radiation from the bombardment of lithium by protons.

* Physics Department, University of Melbourne.

II. PRINCIPLE OF EXPERIMENTS

Consider a stationary nucleus of mass $(A-1)M$ being struck by, and capturing, a proton of mass M and kinetic energy E_p . The compound nucleus will recoil with velocity $\sqrt{(2E_p/M)/A}$. It is assumed to be travelling with this full recoil velocity when it is de-excited to its ground state by the emission of a photon which emerges at angle θ to the recoil direction. The energy of the photon in the laboratory is

$$E = \left(Q + E_p \frac{A-1}{A} \right) \left(1 - \left[Q + E_p \frac{A-1}{A} \right] / 2MAc^2 \right) \left(1 + \frac{\cos \theta}{Ac} \sqrt{\frac{2E_p}{M}} \right), \quad \dots (1)$$

where Q is the binding energy of the proton in the compound nucleus and c the velocity of light. $E_p(A-1)/A$ is the fraction of the proton kinetic energy which is available for excitation in the centre-of-mass system. The second factor is the correction due to the small amount of energy taken by the compound nucleus in recoiling from the emission of the photon, and is actually an expansion terminated at the first order, the approximation being valid for $[Q + E_p(A-1)/A] \ll 2MAc^2$. The third factor is the Doppler correction due to motion of the centre of mass.

The photon energy can therefore be controlled by varying either the proton energy E_p or the angle θ . In these experiments, θ was held constant and E_p varied. Assuming the ground state of the final nucleus to be sharp, the homogeneity in any direction is limited by the energy spread of the proton beam and the thickness of the target, and so can be of the order of 1 keV, although in the present work thicker targets were necessary to obtain sufficient intensity.

The ${}^7\text{Li}(p, \gamma){}^8\text{Be}$ reaction has a Q -value of 17.242 MeV (Ajzenberg and Lauritsen 1955). The ground state of ${}^8\text{Be}$ has a width of a few electron-volts, which is small enough to be ignored, and so the photons from transitions to the ground state have energy given by (1). Transitions also occur to a broad (~ 2 MeV) state at excitation 3 MeV, but since the spread in the resulting photon energies is large compared with the range of variation available they could not produce fine structure effects. The cross section of the ${}^7\text{Li}(p, \gamma){}^8\text{Be}$ reaction exhibits the well-known resonance at 441 keV proton energy, above which it drops to a low but nearly constant level.

In these experiments, the samples of oxygen and copper were exposed to the γ -radiation from the bombardment of thin lithium targets by the proton beam from a 700 keV electrostatic generator. The angle θ was held constant, and the proton energy E_p varied over a range of about 300 keV. The activation induced by photodisintegration was measured at each proton energy.

III. ${}^{16}\text{O}(\gamma, n){}^{15}\text{O}$ EXPERIMENT

Because of the small cross section of this reaction and the low intensity of γ -radiation available, it was necessary to detect the ${}^{15}\text{O}$ produced with the highest possible efficiency. This was done by counting the 1.68 MeV positrons emitted by ${}^{15}\text{O}$, the half-life of which is 118 sec (Ajzenberg and Lauritsen 1955). 4π geometry was obtained by making a scintillation counter in which the scintillator itself was rich in oxygen. The scintillator used was a saturated

solution of *p*-terphenyl in *p*-dioxane, which contains 36 per cent. oxygen by weight, with a trace of diphenylhexatriene to improve the pulse height. It was contained in a small Pyrex vessel in the form of a truncated cone, of semi-angle 30° , with a volume of about 30 cm^3 . The vessel was wrapped in aluminium foil and coupled optically to a photomultiplier (E.M.I. type 6260) with mineral oil.

The pulses from the photomultiplier were fed through a unity gain pre-amplifier and a 1008 amplifier to two 1009 scaling units switched together with connected inputs. The discriminator levels were set to correspond to ionization in the scintillator of 0.6 and 2.2 MeV respectively, so that the difference between their readings gave the counts in a channel bounded by these values.

The thin lithium targets were prepared by evaporation of lithium metal *in vacuo* on to copper disks, and as this was done in a separate vacuum system it was necessary to transfer the targets without exposure to air. This transfer was carried out under an atmosphere of dry nitrogen, so that an unknown amount of lithium would have combined to form Li_3N , but this is still a satisfactory target material.

The proton beam currents used were between 30 and 50 μA . The dioxane scintillator was placed with its axis along the line of the proton beam, and the apex of the truncated cone at the target spot. Thus the γ -radiation entering the dioxane had values of θ ranging from 0 to 30° .

After each bombardment, it was necessary to remove the scintillator from the vicinity of the target, to avoid counts due to stray activities. The counter unit, consisting of scintillator, photomultiplier, preamplifier, and light-tight box, was placed during irradiation in a kinematic mounting with assured reproducibility of position and orientation. When each bombardment was completed, the unit was lifted off the mounting and slid into a "burrow" in a pile of lead bricks. The scintillator was there shielded from stray radiation by at least 10 cm of lead in all directions except at the mouth of the burrow. In order to eliminate counts due to energetic cosmic ray particles, a tray of 17 Geiger counters was inserted in the lead pile just above the scintillator, with a sensitive area about 10 in. square. The dioxane pulses, after passing through a 3 μsec delay line, were put through a gate which was closed for 8 μsec after a pulse from any of the Geiger tubes. Thus the background in the channel was reduced to 4 counts per minute.

The photon flux was monitored by a sodium iodide scintillation counter. The cylindrical NaI(Tl) crystal, $1\frac{1}{2}$ in. long and $1\frac{1}{2}$ in. in diameter, in conjunction with a 6260 photomultiplier, gave pulses which were amplified and fed to a single-channel pulse height analyser. In order that the counter should respond only to the γ -radiation from transitions to the ground state of ^8Be , the channel was set at a level corresponding to an ionization in the crystal of 16 MeV (Campbell and Boyle 1954). The crystal was placed on a line from the target 20° from the forward direction, which was about the mean of the angular range covered by the dioxane.

The proton beam falling on the target was monitored by two circuits in series, a conventional pulse integrator and a "leaky" integrator of the type

first discussed by Snowden (1950). The purpose of the latter was to measure the apparent integrated current as registered by a system with a decaying memory, the decay time of this being equal to that of the activation being studied.

The procedure in each bombardment was to set the proton beam on the target for about 5 min, with the scintillator in position on its kinematic mounting. During the bombardment, the γ -monitor and the pulse integrator were switched on together for a period sufficient to determine the ratio (integrator count)/(γ -count) to sufficient statistical accuracy. At the end of the bombardment, the leaky integrator was read, and simultaneously the stopclock was started which timed the activation counting. After a 30-sec delay, during which the dioxane counter was transferred to the burrow, the counts were recorded during 4 min, followed by a 4-min delay and a further 4-min count. The difference between these two counts was taken as the measure of the activation and the final measure of the photodisintegration cross section at the energy of the bombardment was then

$$\frac{(\text{activation count}) (\text{integrator count})}{(\text{leaky integrator reading}) (\gamma\text{-count})}.$$

No attempt was made to measure the cross section absolutely.

To determine the mean energy of the photons causing the disintegration, equation (1) was used. The mean of E_p was taken to be the incident proton energy less half the target thickness. This is not quite true, because of the variation of proton yield with energy, and if necessary a more accurate mean could be calculated in each case. The mean of $\cos \theta$, over the material of a cone of semi-angle 30° , is 0.933.

Figure 1 shows one of the sets of results obtained. Also shown are points taken in a control experiment when the dioxane scintillator was replaced in the vessel by a solution of *p*-terphenyl in toluene, in which case the only activation detected would be that due to the oxygen in the glass, which would be expected to give between 15 and 20 per cent. of the counts. The results are quite consistent with this estimate.

The vertical lines through the points in Figure 1 denote the 67 per cent. statistical confidence limits. These are shortest in the region of the resonance in the lithium proton capture cross section, where the γ -flux is most intense.

In the case of the results shown in Figure 1, the lithium target used was about 80 keV thick. Several other sets of results were taken, with thinner lithium targets and hence poorer statistical accuracy. In all of these, the points were consistent with a constant photodisintegration cross section* in the region available.

Using the result of Wäffler and Younis (1949) that this cross section has the value 0.54 ± 0.14 mbarn for the lithium resonance radiation, it is possible to place an upper limit on any (γ , n) resonance in this region. It can be concluded

* In a note on a preliminary experiment (Campbell 1954) it was claimed that a resonance had been detected in this range. This was erroneous, and resulted from stray target activations, mainly ^{13}N .

that a sharp resonance between 17.6 and 17.9 MeV would certainly have been observed if its integrated cross section were as great as 0.05 MeV-mbarn.

This result is to be compared with the results of the betatron work mentioned above. Katz *et al.* (1954) covered this range of photon energies and detected no resonance. Penfold and Spicer (personal communication), however, report a break corresponding to an energy level in ^{16}O at 17.71 MeV with integrated cross section 0.18 MeV-mbarn.

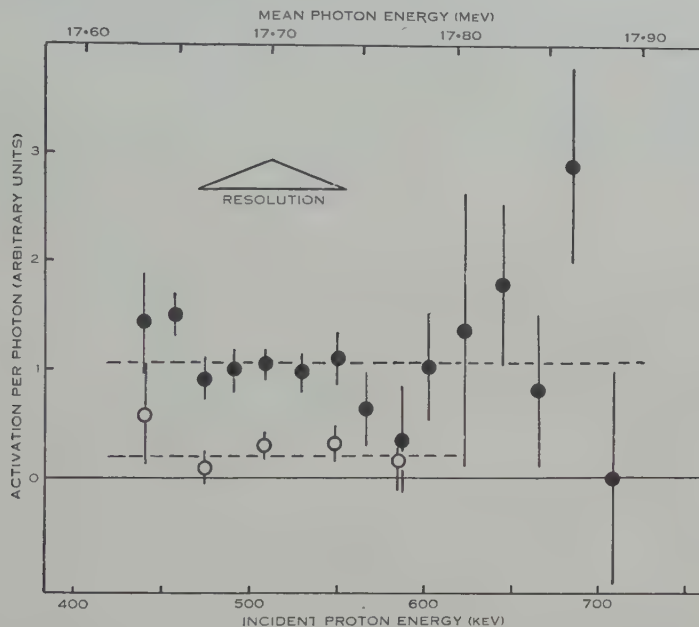


Fig. 1.—A set of results for $^{16}\text{O}(\gamma, n)^{15}\text{O}$. Closed circles: dioxane scintillator; open circles: toluene scintillator.

IV. $^{63}\text{Cu}(\gamma, n)^{62}\text{Cu}$ EXPERIMENT

In this experiment, the copper metal sample, which was in the form of an annulus of the exterior of a cone, 2 in. in diameter with a central hole $\frac{3}{8}$ in. in diameter, was placed around the target so that it intercepted radiation emerging at 90° . The Doppler spread in the radiation entering the sample depended on the angle it subtended at the target. Two samples were used, one subtending $82^\circ \leq \theta \leq 98^\circ$ and the other $60^\circ \leq \theta \leq 98^\circ$.

After irradiation, the sample was removed to between a pair of sodium iodide scintillation counters in coincidence, which detected the positron annihilation radiation from ^{62}Cu . Two flat aluminium cones were sprung together to form a sandwich around the sample, ensuring that all positrons were absorbed. The counting system was in a pile of lead bricks to reduce background.

The procedure was similar to that in the oxygen experiment, the leaky integrator now being set to a half-life of 10.0 ± 0.1 min, the mean of recent

determinations of the half-life of ^{62}Cu (U.S. National Bureau of Standards 1950). The sodium iodide counter measuring the flux of ground-state γ -radiation was placed below the target, at the mean angle subtended by the sample.

The activation was counted for 20 min, followed by a 20-min delay and a further 20-min count. The relative photodisintegration cross section was found as before.

The results with the thinner sample indicate a dip in the cross section at about 550 keV incident proton energy, but it is of the same magnitude as the statistical uncertainties. The results for the thicker sample are shown in Figure 2, and a similar dip appears here too.

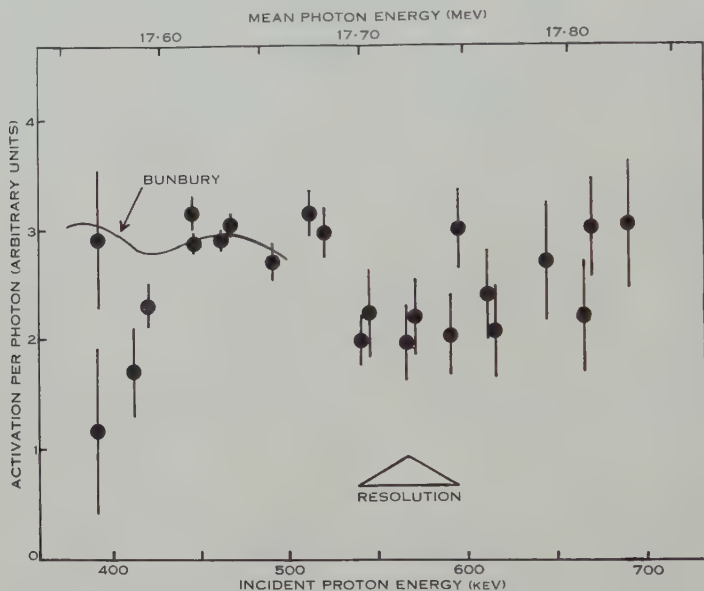


Fig. 2.—A set of results for $^{63}\text{Cu}(\gamma, n)^{62}\text{Cu}$. Also shown are results of Bunbury (1954).

While this experiment was being planned, Bunbury (1954) reported an experiment on this reaction using the same lithium γ -radiation. In that work the radiation from the 441 keV proton capture resonance was used, the energy being varied by alteration of angle. This has the advantage of high γ -intensity, but the range of energy variation is more limited. The curve obtained by Bunbury is shown on Figure 2, with ordinate adjusted to fit. The statistical uncertainties in the present experiment would have obscured the fine structure observed by Bunbury, and the dip at 550 keV is beyond the range of photon energy obtainable by Doppler shift of the resonance radiation.

It is necessary to interpret the copper results in terms of known theory. There are little data available about medium-weight nuclei at high excitation, but Blatt and Weisskopf (1952, pp. 371-2) have given a rough, semi-empirical formula based on a statistical nuclear model. With the constants given for mass number 63, this formula gives the mean spacing of all levels to be about

24 eV at excitation 17.6 MeV. An experiment such as the present one could not resolve these levels individually, but the average integrated cross section within the finite interval of resolution would be expected to show some variation, just as the needle of a count-rate meter fluctuates even when its integrating time is long compared with the mean pulse spacing.

Since each point in Figure 2 has a total resolution width of about 50 keV, then each would include the effects of about 2000 levels. If these levels are spaced randomly along the energy scale, then the number of levels affecting each point is statistically distributed about 2000 according to the Poisson distribution, with standard deviation $\sqrt{2000}$. Assuming in the first instance that all levels have the same integrated cross section, then the apparent cross section of independent points will fluctuate with proportional standard deviation 2 per cent. To take account of the differences in level cross sections, it might be assumed that one-quarter of them are large and equal (these would be the ones giving electric dipole transitions) and the remainder small enough to be ignored. The result is now an expected variation in apparent cross section with standard deviation 4 per cent. This is to be compared with the experimental standard deviation of about 15 per cent., which would be the fluctuation of a smooth line of best fit through the points of Figure 2. These figures, while not being in close agreement, are within the same order of magnitude, and taking account of the rudimentary nature of the level density formula it can be concluded that the results are not inconsistent with this model.

V. ACKNOWLEDGMENTS

The author has appreciated the continued encouragement of Professor L. H. Martin, and acknowledges the assistance of Mr. B. M. Spicer in communication of results prior to publication and in discussion.

VI. REFERENCES

- AJZENBERG, F., and LAURITSEN, T. (1955).—*Rev. Mod. Phys.* **27**: 77.
 BLATT, J. M., and WEISSKOPF, V. F. (1952).—"Theoretical Nuclear Physics." (John Wiley & Sons: New York.)
 BUNBURY, D. St. P. (1954).—*Proc. Phys. Soc. Lond. A* **67**: 1106.
 CAMPBELL, J. G. (1954).—*Phys. Rev.* **95**: 1357.
 CAMPBELL, J. G., and BOYLE, A. J. F. (1954).—*Aust. J. Phys.* **7**: 284.
 DAVEY, W. G., and MOON, P. B. (1953).—*Proc. Phys. Soc. Lond. A* **66**: 956.
 ILAKOVAC, K. (1954).—*Proc. Phys. Soc. Lond. A* **67**: 601.
 KATZ, L., HASLAM, R. N. H., HORSLEY, R. J., CAMERON, A. G. W., and MONTALBETTI, R. (1954).—*Phys. Rev.* **95**: 464.
 KUHN, W. (1929).—*Phil. Mag.* **8**: 812.
 MALMFORS, K. G. (1952).—*Ark. Fys.* **6**: 49.
 MOON, P. B. (1951).—*Proc. Phys. Soc. Lond. A* **64**: 76.
 PHILLIPS, K. (1953).—*Phil. Mag.* **44**: 533.
 SNOWDON, S. C. (1950).—*Phys. Rev.* **78**: 299.
 U.S. NATIONAL BUREAU OF STANDARDS (1950).—Nuclear data. Circ. U.S. Nat. Bur. Stand. No. 499.
 WÄFFLER, H., and YOUNIS, S. (1949).—*Helv. Phys. Acta* **22**: 614.

PHOTOPROTONS FROM FLUORINE

By W. B. LASICH,* E. G. MUIRHEAD,† and G. G. SHUTE†

[Manuscript received August 31, 1955]

Summary

The energy and angular distributions of protons produced in the $^{19}\text{F}(\gamma, p)$ reaction have been obtained for the photon energy range 10–16.5 MeV. An integrated cross section of 18 MeV-mbarn was obtained for this region. The angular distribution was found to be of the form $(a+b \sin^2 \theta)$ with the ratio b/a strongly energy dependent. The higher energy protons show evidence of an independent particle-type interaction which is in plausible agreement with the shell model structure of fluorine.

The range-energy relation for the target gas, CF_4 , is presented in Appendix I.

I. INTRODUCTION

Recent experiments on photonuclear reactions have been directed towards elucidating the mechanisms of photon absorption by nuclei undergoing photodisintegration. Photoneutron production in light nuclei has been studied by Katz *et al.* (1954) and Penfold and Spicer (1955) who showed that, for the (γ, n) reactions studied, photons excite the nucleus predominantly by level absorption with the formation of a compound nucleus state. The absorption is probably electric quadrupole or magnetic dipole in character at excitation energies immediately above the binding energy of the nucleon being ejected, changing to electric dipole as the region of the “giant resonance” is approached. The relative importance of the compound nucleus formation and the alternative direct or independent-particle interaction (Courant 1951, Wilkinson 1955) has yet to be determined, especially for light nuclei. In the case of carbon the protons ejected at the giant resonance at 22 MeV were found to have an angular distribution as predicted by the independent-particle model (Mann, Stephens, and Wilkinson 1955).

Experiments on the photodisintegration of light nuclei are particularly suitable for the study of photon absorption, since the level density in the excited nucleus is still small enough to permit the resolution of products coming from discrete energy levels. Much of the experimental work relevant to this problem has been summarized by Titterton (1955).

The present experiment was undertaken to study the production of photoprotons from a light nucleus. Fluorine was chosen because of its low proton binding energy and also because it is free of complications arising from the presence of more than one isotope. A gas target was considered preferable for the measurement of angular distributions, and a suitable gas, tetrafluoromethane

* Australian Atomic Energy Commission, University of Melbourne.

† Physics Department, University of Melbourne.

(CF₄), was found to be available commercially with a high degree of purity. The energy and angular distributions of protons from fluorine were obtained for the photon region 10–16.5 MeV using the X-ray beam from the Melbourne synchrotron and a reaction chamber containing nuclear emulsions.

II. EXPERIMENTAL

The X-rays were produced from a conventional betatron target, a 0.005 in. platinum flag. The beam was reduced to an angular width of 0.022 radian by a lead collimator, and then passed through a 0.001 in. aluminium window into the reaction chamber. Here the mean beam diameter was 1.7 cm as determined by radiographs taken on the emergent side of the chamber. The reaction chamber, based on the design of Fuller (1950), was used previously by Spicer (1953) for the ¹⁴N(γ , p) reaction. The gas target was maintained at a pressure of half an atmosphere. Two 100 μ C2 emulsions were placed relative to the beam as shown in Figure 1. The data were taken from two runs, each

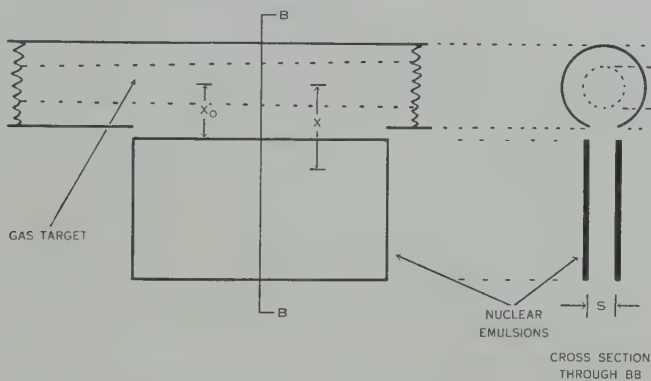


Fig. 1.—Positioning of nuclear emulsions relative to X-ray beam. Mean beam diameter, $D=1.7$ cm; emulsion spacing, $S=1.3$ cm; beam centre to emulsion, $X_0=2.3$ cm; beam centre to average swathe, $X=3.7$ cm.

target irradiation being approximately 1100 r. The chamber was surrounded by 20 cm of paraffin and a similar thickness was placed between the doughnut and the collimator to reduce the neutron background. However, the X-ray beam was allowed to pass through a 1.5 cm hole in the paraffin so as to preserve more closely the "thin" target spectral distribution. The exposure was obtained from the readings of a 25 r Victoreen thimble contained in an 8 cm "Perspex" cube, the latter being placed in the beam opposite the exit end of the reaction chamber. The photon flux through the chamber was deduced from this "Lucite-roentgen" reading using the tables of Katz and Cameron (1951).

The synchrotron energy scale was calibrated from the ratio of induced activities in copper and silver using the activation curves of Diven and Almy (1950), and allowed the peak energy of the X-ray distribution to be determined as 16.5 ± 0.5 MeV. Track measurements on the emulsions were made on Cooke,

Troughton, and Simms model 4000 microscopes, a $45\times$ objective with $8\times$ eyepiece giving a convenient size field of some $200\ \mu$ diameter. Tracks were accepted from swathes of $160\ \mu$ width so as to avoid possible bias in the angular distributions due to loss of tracks commencing near the periphery of the field of view. The usual angle, range, and dip measurements were recorded for each track. The criterion for acceptance of a track was that it should appear to cross the emulsion surface and have a direction and dip compatible with an origin in the irradiated part of the gas target. Using selected areas on the emulsions, tracks were accepted for angles θ , in the range $30\text{--}150^\circ$ with a dip angle, δ° , satisfying the condition,

$$\delta \leq 24 \sin \theta (3.5/X), \quad \dots\dots\dots (1)$$

where X is the beam-swathe distance in cm as defined in Figure 1, and θ is the angle between the incident photon and the proton.

The range of each proton accepted as coming from the target was determined from the known geometrical set-up of the emulsions relative to the beam. To obtain each proton energy, the range-energy relations used were (i) the tables of Wilkins (1951) for C2 emulsions, and (ii) the table for CF_4 presented in Appendix I (Table 4). In practice, it was found convenient to convert the residual range in the emulsion to the CF_4 gas range of a proton with equivalent energy, which, when added to range in the gas target, permitted the tracks for the energy distribution to be grouped according to the total equivalent gas range.

The main uncertainty in the determination of proton energies was due to the finite thickness of the gas target used. In accumulating tracks for the energy and angular distributions, different criteria were required in order to preserve good energy resolution in the one case and adequate angular coverage in the other. Since the effective thickness of the target depended on $\text{cosec } \theta$, the data for the energy spectrum were restricted to the range of θ from 45 to 135° . The uncertainty in proton energy arising from the target thickness varied from 7 per cent. for 2 MeV protons to 1.7 per cent. for 6 MeV tracks. In addition, proton range straggling amounted to 2–3 per cent. The resulting overall uncertainty in energy ranged from ~ 0.15 MeV for 2 MeV protons, to ~ 0.12 MeV for 6 MeV protons. Hence it was not possible to demonstrate the existence of distinct proton groups with energies closer than approximately 0.3 MeV.

For the angular distribution, on the other hand, the lower limit for the acceptance of tracks was set in principle by the minimal residual range in the emulsion which would be easily identified by an observer. This was set at a range of $12\ \mu$ (0.9 MeV), which, together with the angular limits used ($30\text{--}150^\circ$), resulted in a lower limit for the proton energy of 2.5 MeV. The angular distribution, plotted for 15° intervals of θ , was obtained from the relation between the relative differential cross section $d\sigma$, and the number of tracks $N(\theta)$ per 15° interval (Fuller 1950) :

$$N(\theta) = \text{const. } d\sigma \overline{\sin \theta}, \quad \dots\dots\dots (2)$$

with $\overline{\sin \theta}$ the corresponding mean value of $\sin \theta$ for each 15° interval.

Background tracks arose from two possible sources : (i) those generated in the beam region due to reactions other than the $^{19}\text{F}(\gamma, p)$ reaction, and (ii) those produced in the emulsions as neutron-proton recoils.

The thresholds for the majority of type (i) reactions are given in Table 1. The thresholds for (n, p) reactions in carbon and fluorine are such that the number of protons produced which would reach the emulsions was negligible. In fluorine, for example, using an estimate of the flux and energy distribution of fast neutrons from an X-ray target (Wäffler 1954), it was shown that the number of protons produced in the (n, p) reaction with energy sufficient to reach the emulsions was of the order 10^{-5} times the expected number of photoprotons.

TABLE 1
THRESHOLDS FOR POSSIBLE COMPETING REACTIONS

Reaction	Threshold (MeV)	Reference
$^{19}\text{F}(\gamma, p)$	7.97	Drummond (1955)
$^{19}\text{F}(\gamma, d)$	13.80	"
$^{19}\text{F}(\gamma, t)$	11.69	"
$^{19}\text{F}(\gamma, \alpha)$	4.01	"
$^{19}\text{F}(\gamma, np)$	16.02	"
$^{12}\text{C}(\gamma, p)$	15.94	"
$^{12}\text{C}(\gamma, d)$	25.17	"
$^{12}\text{C}(\gamma, \alpha)$	7.36	"
$^{19}\text{F}(n, p)$	3.72	Ajzenberg and Lauritsen (1955)
$^{12}\text{C}(n, p)$	12.6	" "

Of the other reactions, photoprotons produced in the carbon failed to reach the emulsions, for the peak irradiation energy of 16.5 MeV chosen for this experiment. Likewise, the stopping power of the gas was sufficient to prevent the recording of deuterons and tritons ejected from fluorine. Preliminary runs showed that a significant number of tracks recorded came from the brass walls of the reaction chamber ; accordingly, the latter was lined completely with graphite walls $\frac{1}{8}$ in. thick.

An estimate of the background tracks of type (ii) crossing the surface was made during the course of scanning. The number of tracks in the range $\theta=210-330^\circ$ was recorded. It was not possible to identify the direction of individual tracks but the majority was shown by the following method to have started in the emulsion. The average volume density of tracks was determined from sample counts made at various depths in the emulsion. Using this, an estimate of the number of tracks leaving the emulsion was made and this agreed with the number observed at the upper surface. The contribution of these protons to the total number of tracks observed in the acceptance sector ($\theta=30-150^\circ$) was calculated by assuming the distribution of the background to be symmetrical about a direction parallel to the beam axis. The distribution so obtained was consistent with an independent estimate obtained by scanning

a sample area ($=0.36 \text{ cm}^2$) and recording all tracks appearing to cross the emulsion surface with $\theta=30\text{--}150^\circ$. A chart was prepared showing the distribution of tracks with large angles of dip ($\delta=30\text{--}60^\circ$) and all possible values of θ . From the space distribution in δ the number of tracks occurring in the allowed dip region given by equation (1) was deduced. This estimate agreed within 10 per cent. with the former estimate, i.e. within the statistical error associated with a sample of 156 tracks.

In order to subtract the background, the same set of operations was performed on the background sample as for the main group of tracks. A fictitious energy was assigned to each background track, i.e. the energy which would be possessed by a photoproton of similar range and dip and with angle equal to $360^\circ-\theta$. The same dip criterion (equation (1)) was also applied so as to have the same solid angle weighting factor for both sets of data. The backgrounds determined in this way were subtracted from each of the energy and angular distributions, as given in Figures 2 and 3.

III. RESULTS

(a) *Energy Distribution*

Figure 2 (a) shows the number of photoprotons with energy greater than 2 MeV obtained from an emulsion area of 4.8 cm^2 after the subtraction of background. The latter is presented for reference in Figure 2 (b). The limit of resolution for discrete proton groups is about 0.3 MeV. For reference, the positions of every proton energy group expected from the known level structure in ^{19}F and ^{18}O are indicated by arrows in Figure 2 (a). The poor resolution and statistics prevent any correlation between these positions and the observed data to be established.

(b) *Angular Distribution*

The angular distributions for the two proton energy groups, 4.4–8.5 MeV and 2.9–4.4 MeV, are presented in Figures 3 (a) and 3 (b) respectively. The division at 4.4 MeV was made to obtain a sufficiently large sample of higher energy protons for an angular distribution. The lower limit of 2.9 MeV was taken in order to reduce the effect of background. In this way the percentage backgrounds for the higher and lower energy groups were 10 and $13\frac{1}{2}$ respectively. The simplest theoretical form of the angular distribution for the data is of the form $a+b\sin^2\theta$, since there is no apparent forward asymmetry in the data. A least squares fit for each group resulted in the smooth curves shown in Figure 3. The higher energy group can best be fitted, within the statistical uncertainties, by a $\sin^2\theta$ distribution, while the lower energy group tends more towards isotropy.

(c) *Integrated Cross Section*

It is not possible to obtain a precise value for the integrated cross section for the $^{19}\text{F}(\gamma, p)^{18}\text{O}$ reaction from the proton energy distribution without more detailed knowledge of the levels in ^{18}O . At present there is evidence for levels at 1.98 MeV and possibly at 2.42 MeV (Ajzenberg and Lauritsen 1955).

However, a lower limit can be derived from Figure 2 by assuming that all proton groups were due to transitions to the ^{18}O ground state only. The energy of each corresponding photon can then be determined from the relation

$$E = \frac{19}{18}E_p + 7.95, \quad \dots\dots\dots (3)$$

and the integrated cross section is obtained from the relation

$$N(E_0) = k \overline{P(E, E_0)} \int_{10}^{16.5} \sigma(E) dE, \quad \dots\dots\dots (4)$$

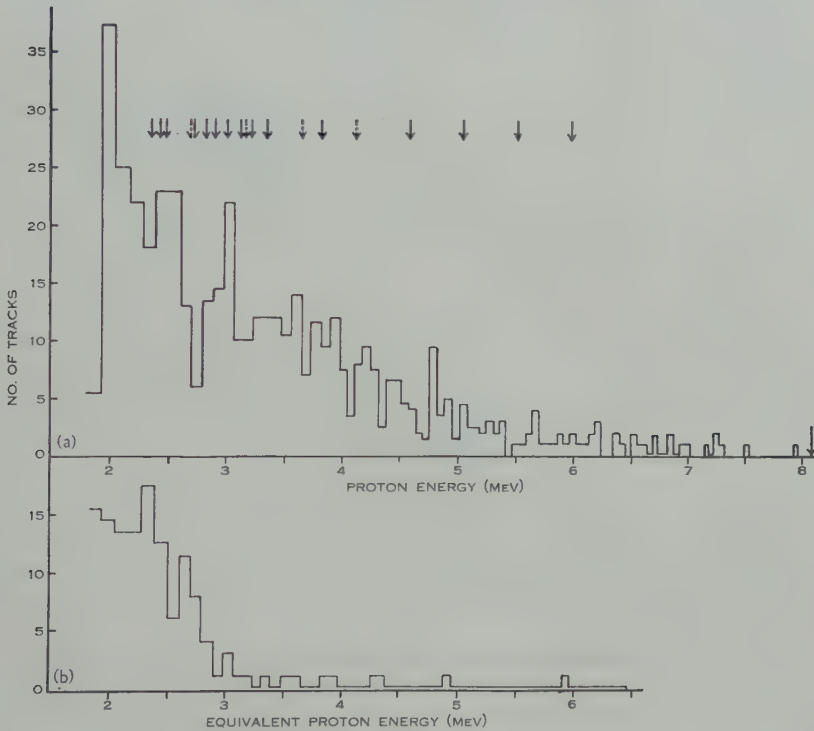


Fig. 2 (a).—Energy distribution of photoprotons from ^{19}F . The expected positions of proton groups between known energy levels in ^{19}F and ^{18}O are indicated by arrows above the histogram as follows:

- ↓ $^{19}\text{F}^*$ to ^{18}O (ground state),
- ↓ $^{19}\text{F}^*$ to $^{18}\text{O}^*$ (1.98 MeV state).

(The energy levels in ^{19}F were observed in $^{18}\text{O}(p,n)^{18}\text{F}$ (Ajzenberg and Lauritsen 1955).) The limit of resolution is 0.3 MeV.

Fig. 2 (b).—Corresponding background tracks.

where $N(E_0)$ is the number of photoprotons in Figure 3 and $\overline{P(E, E_0)}$ is the average value of the photon flux per $\frac{1}{2}$ MeV interval between 10 and 16.5 MeV as taken from the tables of Katz and Cameron (1951). The value of k ($=1.86 \times 10^{21}$) for the present experiment was obtained from the geometry of the chamber, the operating conditions, and the assumed angular distributions.

for the groups in Figure 3. The result so obtained was 6 MeV-mbarn. This value is unrealistic owing to the probable importance of the 1.98 MeV level with $J=2$ in ^{18}O . In a similar way, an upper limit to the integrated cross section, assuming that where energetically possible all protons involved transitions to this state, leads to the value of 19 MeV-mbarn. Making the crude assumption that the relative importance of the ground and first excited states in ^{18}O is the ratio of their statistical weights, i.e. 1 : 5, gives a provisional estimate of the integrated cross section as 18 MeV-mbarn for the photon energy region 10–16.5 MeV.

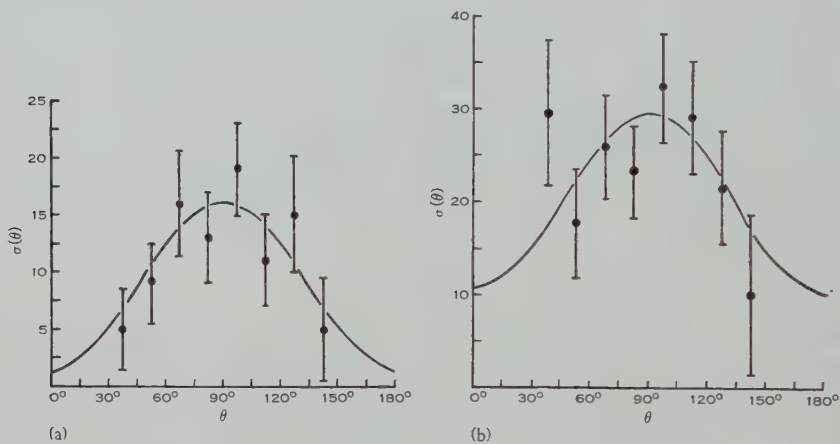


Fig. 3.—Angular distribution of photoprotons from ^{19}F . θ =angle between incident photon and the proton, $\sigma(\theta)$ =differential cross section (arbitrary units). Vertical lines represent standard deviations of the points.

(a) Energy range 4.4–8.5 MeV. No. of tracks=75, percentage background=10, least squares fit to data is shown by curve $(1 \pm 5) + (15 \pm 7) \sin^2 \theta$.

(b) Energy range 2.9–4.4 MeV. No. of tracks=180, percentage background=13½, least squares fit to data is shown by curve $(10 \pm 8) + (19 \pm 11) \sin^2 \theta$.

For the reasons mentioned above, it is not possible to estimate the variation of cross section within the region studied. However, making the same assumption as for the integrated cross section, it can be shown that the cross section decreases monotonically with energy by a factor of two within the range under consideration.

IV. INTERPRETATION

The angular distributions of photoprotons presented in Figure 3 can be represented by the expression $(a + b \sin^2 \theta)$ with a and b positive or zero. It is of interest to compare the experimental distributions with those predicted for the process involving the formation of compound nucleus states in ^{19}F with assumed values of spin and parity. These are presented in Tables 2 and 3 for the ground and first excited states of ^{18}O . For the lower energy group (Fig. 3 (b)), the experimental angular distribution is more nearly consistent with E1 or M1 absorption with transitions to the ^{18}O ground state (Table 2), depending on whether the levels available in ^{19}F with spin 3/2 have negative or positive parity in the energy region considered. It does not seem plausible to fit these data

TABLE 2
THEORETICAL ANGULAR DISTRIBUTION FOR $^{19}\text{F}(\gamma, p)^{18}\text{O}$ REACTION FOR TRANSITIONS TO
THE GROUND STATE ONLY
 $I = \frac{1}{2} +$ throughout†

Type of Photon Absorption	L	J	l'	s'	$\sigma(\theta)$
E1	1	$\frac{1}{2}-$	1	$\frac{1}{2}+$	isotropic
E1	1	$3/2-$	1	$\frac{1}{2}+$	$2 + 3s^2$
M1	1	$\frac{1}{2}+$	0	$\frac{1}{2}+$	isotropic
M1	1	$3/2+$	2	$\frac{1}{2}+$	$2 + 3s^2$
E2	2	$3/2+$	2	$\frac{1}{2}+$	$1 + c^2$
E2	2	$5/2+$	2	$\frac{1}{2}+$	$1 + 6c^2 - 5c^4$

† Here I =spin of initial state of ^{19}F ($=\frac{1}{2}+$),
 J =spin of the excited state, $^{18}\text{F}^*$,
 L =angular momentum of incoming photon,
 l' =angular momentum of emitted proton,
 s' =final channel spin,
 $c = \cos \theta$,
 $s = \sin \theta$.

TABLE 3
THEORETICAL ANGULAR DISTRIBUTIONS FOR $^{19}\text{F}(\gamma, p)^{18}\text{O}$ REACTION FOR TRANSITIONS TO THE FIRST
EXCITED STATE OF ^{18}O AT 1.98 MeV
 $I = \frac{1}{2} +$ throughout*

Type of Absorption	L	J	l'	s'	$\sigma(\theta)$
E1	1	$\frac{1}{2}-$	1	$3/2+$	Isotropic
	1	$\frac{1}{2}-$	3	$5/2+$	Isotropic
	1	$3/2-$	1	$3/2+$	$4 + 3c^2$
	1	$3/2-$	1	$5/2+$	$21 - 3c^2$
	1	$3/2-$	3	$5/2+$	$19 + 3c^2$
M1	1	$\frac{1}{2}+$	2	$3/2+$	Isotropic
	1	$\frac{1}{2}+$	2	$5/2+$	Isotropic
	1	$3/2+$	0	$3/2+$	Isotropic
	1	$3/2+$	2	$5/2+$	$23 + 15c^2$
	1	$3/2+$	4	$5/2+$	$33 - 15c^2$
E2	2	$3/2+$	0	$3/2+$	Isotropic
	2	$3/2+$	2	$3/2+$	Isotropic (accidentally)
	2	$3/2+$	2	$5/2+$	$33 - 15c^2$
	2	$3/2+$	4	$5/2+$	$23 + 15c^2$
	2	$5/2+$	2	$3/2+$	$56 - 105c^2 + 140c^4$
	2	$5/2+$	4	$3/2+$	$119 + 420c^2 - 315c^4$
	2	$5/2+$	0	$5/2+$	Isotropic
	2	$5/2+$	2	$5/2+$	$189 + 210c^2 - 315c^4$
	2	$5/2+$	4	$5/2+$	$203 - 210c^2 + 315c^4$

* See footnote to Table 2 for explanation of symbols.

with the predicted distributions for electric quadrupole absorption (Table 2) or from any of the channels listed in Table 3, unless one regards the experimental distribution to be either (i) wholly isotropic or (ii) made up of isotropic distributions contributed from various channels available in the tables plus a $\sin^2 \theta$ component superimposed. Considering the higher energy group (Fig. 3 (a)), the distribution appears at variance with the predictions for the compound nucleus states listed in Tables 2 and 3. The distribution suggests a direct photoelectric interaction. The predicted angular distributions for the latter in the case of E1 absorption by a proton with initial angular momentum in the nucleus of l is :

$$\text{for transitions } l \text{ to } l+1 : \sigma(\theta) = 2l + (l+2) \sin^2 \theta, \quad \dots\dots (5)$$

$$\text{for transitions } l \text{ to } l-1 : \sigma(\theta) = 2(l+1) + (l-1) \sin^2 \theta. \quad \dots (6)$$

Both types of transition are predicted by the theory of Courant (1951), while that of Wilkinson (1955) allows the " l to $l+1$ " transition only. Thus the ejection of an " s " proton from the fluorine nucleus on these theories would lead to a $\sin^2 \theta$ distribution in agreement with Figure 3 (a). It is significant that the ground state proton in ^{19}F has an s -wave designation on current nuclear shell theory.

Our results for the angular distributions can be compared with those of neighbouring light nuclei for the photon region below the giant resonance. On the one hand, the distributions for aluminium (Diven and Almy 1950) and magnesium (Toms and Stephens 1951) were found to be isotropic, or nearly so. On the other, the angular distribution for oxygen (Spicer 1955) was consistent with predictions involving compound nucleus formation, i.e. for the photon range 13.5–18.7 MeV, the distribution was consistent with E2 or M1 photon absorption. The difference in behaviour between oxygen and fluorine may well be due to the fact that the former is a closed shell nucleus while fluorine is a "closed-shell-plus-one" nucleus. The angular distribution of photoprotons from carbon at the giant resonance has been reported by Mann, Stephens, and Wilkinson (1955) who found their data was fitted best by predictions based on the independent particle character of the proton being ejected.

There is little to be gained from a detailed examination of the energy distribution of protons given in Figure 2. Owing to the density of levels and the inferior resolution, no conclusion can be drawn from it regarding the relative amounts of level absorption and continuous absorption of photons as required by the compound nucleus and independent particle interactions. The integrated cross section estimate of 18 MeV-mbarn for the photon range 10–16.5 MeV is comparable to that of the $^{19}\text{F}(\gamma, n)$ reaction in the same region, namely, 12 MeV-mbarn for the range 10.4–17 MeV (Horsley, Haslam, and Johns 1952).

V. CONCLUSIONS

The energy and angular distributions of the photoprotons from fluorine have been studied. The energy distribution does not provide evidence for absorption of photons by levels in ^{19}F . The angular distributions provide more

significant results. The higher proton energy group is best fitted with a $\sin^2 \theta$ distribution which would be interpreted as a direct ejection of an "s" proton from the ^{19}F nucleus. This is consistent with the nuclear shell designation of the odd proton in fluorine. The angular distribution of the lower energy protons is more nearly isotropic and is not inconsistent with the predictions for the formation of intermediate compound nucleus states in ^{19}F . This distribution may be interpreted either as a mixture of direct photonuclear absorption and components from various channels listed in Tables 2 and 3, or, if solely due to compound nucleus formation in this region, as due to electric dipole or magnetic dipole absorption of photons by fluorine. The integrated cross section for the region studied is of the same order as that of the $^{19}\text{F}(\gamma, n)$ reaction for the same region.

VI. ACKNOWLEDGMENTS

The authors wish to thank Professor L. H. Martin for his continued interest and advice during the course of the experiment. Thanks are also due to Dr. J. H. Smith of the University of Illinois and to Mr. B. M. Spicer for access to unpublished work and advice on the interpretation of the angular distributions.

VII. REFERENCES

- AJZENBERG, F., and LAURITSEN, T. (1955).—*Rev. Mod. Phys.* **27** : 77.
- ARON, W. A., HOFFMAN, B. G., and WILLIAMS, F. C. (1949).—Range-energy curves (2nd revision). U.S. Atomic Energy Commission Rep. A.E.C.U. 663.
- COURANT, E. D. (1951).—*Phys. Rev.* **82** : 703.
- DIVEN, B. C., and ALMY, G. M. (1950).—*Phys. Rev.* **80** : 407.
- DRUMMOND, J. E. (1955).—*Phys. Rev.* **97** : 1004.
- FULLER, E. G. (1950).—*Phys. Rev.* **79** : 303.
- HIRSCHFELDER, J. O., and MAGEE, J. L. (1948).—*Phys. Rev.* **73** : 208.
- HORSLEY, R. J., HASLAM, R. N. H., and JOHNS, H. E. (1952).—*Phys. Rev.* **87** : 756.
- KATZ, L., and CAMERON, A. G. W. (1951).—*Canad. J. Phys.* **29** : 518.
- KATZ, L., HASLAM, R. N. H., HORSLEY, R. J., CAMERON, A. G. W., and MONTALBETTI, R. (1954).—*Phys. Rev.* **95** : 464.
- LEVINGER, J. S., and BETHE, H. (1950).—*Phys. Rev.* **78** : 115.
- LINDHARD, J., and SCHARFF, M. (1953).—*Math.-fys. Medd.* **27**, No. 15.
- LIVINGSTON, M. S., and BETHE, H. (1937).—*Rev. Mod. Phys.* **9** : 263.
- MANN, A. K., STEPHENS, W. E., and WILKINSON, D. H. (1955).—*Phys. Rev.* **97** : 1184.
- PENFOLD, A. S., and SPICER, B. M. (1955).—*Phys. Rev.* (in press).
- SPICER, B. M. (1953).—*Aust. J. Phys.* **6** : 391.
- SPICER, B. M. (1955).—*Phys. Rev.* **99** : 33.
- TITTERTON, E. W. (1955).—*Progr. Nucl. Phys.* **4** : 1.
- TOMS, M. E., and STEPHENS, W. E. (1951).—*Phys. Rev.* **82** : 709.
- WÄFFLER, H. (1954).—"The Brown Boveri Betatron." p. 17. (Brown Boveri and Co.: Baden.)
- WILKINS, J. J. (1951).—Range-energy relations for Ilford nuclear emulsions. A.E.R.E. Rep. G/R 664.
- WILKINSON, D. H. (1955).—"Proceedings of the 1954 Glasgow Conference on Nuclear Physics." (Pergamon Press: London.)
- WILSON, R. R. (1941).—*Phys. Rev.* **60** : 749.

APPENDIX I

Range-Energy Relation for Protons in Tetrafluoromethane (CF₄)

The range-energy relation in CF₄ has been calculated for proton energies from 0–10 MeV. For energies greater than 2 MeV, the usual Bethe-Möller formula for differential stopping power was used (Livingston and Bethe 1937). For low energy protons, an empirical treatment of the variation of stopping power adopted was similar to that of Lindhard and Scharff (1953).

Low Energy Region.—The stopping power B is given by the empirical relation

$$B/Z = a_1 x^{\frac{1}{2}} + a_2 x^{3/2}, \quad \dots \dots \dots (\text{A1})$$

where Z = atomic number of the material and $x = (v^2/2\alpha^2 Z)$, v being the proton velocity and α the fine structure constant. The numerical values of a_1 and a_2 are fitted so as to give the experimental values of B/Z for air, and have the values $a_1 = 1.40$ and $a_2 = -0.06$ for 2 MeV protons.

For a molecule with s_i atoms of atomic number Z_i , and stopping power B_i , the effective stopping power

$$B_e = \sum_i s_i B_i. \quad \dots \dots \dots (\text{A2})$$

Now the energy loss of protons in a gas under standard conditions is given by (Hirschfelder and Magee 1948)

$$dE/dR = -0.006094(B/E). \quad \dots \dots \dots (\text{A3})$$

These expressions can be combined to give for the differential stopping power in MeV/cm :

$$dE/dR = -0.006094 \eta a_1 E^{-\frac{1}{2}} (1 + kE) \sum_i (s_i Z_i^{\frac{1}{2}}), \quad \dots \dots (\text{A4})$$

where

$$\eta = \alpha^{-1} (2/Mc^2)^{\frac{1}{2}} = 6.327,$$

and

$$k = \eta^2 (a_2/a_1) (\sum_i s_i Z_i^{-\frac{1}{2}}) / (\sum_i s_i Z_i^{\frac{1}{2}}),$$

$$\alpha = 1/137,$$

$$Mc^2 = 938.2 \text{ MeV}.$$

In the case of CF₄, $k = -0.0551$.

Integrating equation (4) leads to an expression for the range

$$R = \frac{2}{3} \frac{E^{3/2}}{0.5571 a_1} \left(1 - \frac{3}{5} kE + \frac{3}{7} k^2 E^2 \right), \quad \dots \dots (\text{A5})$$

which gives the value $R = 2.59$ cm for $E = 2$ MeV.

High Energy Region.—The differential stopping power in this region was calculated in the usual way from the Bethe-Möller formula

$$\frac{dE}{dR} = -\frac{4\pi e^4 z^2}{mv^2} NZ \left\{ \ln \frac{2mv^2}{I} - \ln (1 - \beta^2) - \beta^2 \right\}, \quad \dots (\text{A6})$$

using a value of $k' = I/Z = 11.5$, as found for air and aluminium by Wilson (1941). The differential stopping powers for carbon were taken from the tables of Aron, Hoffman, and Williams (1949), and the values for fluorine were calculated in a similar way. By means of the usual assumption that the differential stopping powers are additive, the range-energy relation for CF_4 was obtained by numerical integration for protons with energy between 2 and 10 MeV.

Combining this with the lower energy determination, the values of differential stopping power and range are listed in Table 4.

TABLE 4
DIFFERENTIAL STOPPING POWER AND RANGE-ENERGY RELATION FOR
 CF_4 AT 15 °C AND 760 MM HG

Proton Energy (MeV)	Stopping Power relative to Air	Range (cm)
0	2.69	0
0.5	2.70	0.30
1	2.71	0.85
2	2.73	2.59
3	2.76	5.01
4	2.78	8.14
5	2.79	11.93
6	2.81	16.37
7	2.81	21.43
8	2.82	27.10
9	2.82	33.36
10	2.83	40.21

THE DIFFUSION OF SLOW ELECTRONS IN DEUTERIUM

By BARBARA I. H. HALL*

[Manuscript received July 11, 1955]

Summary

The agitational energies and drift velocities of slow electrons diffusing in deuterium are measured as a function of the ratio Z/p of the electric field strength Z to the gas pressure p . The lateral spread of the diffusing electron stream is measured, which enables Townsend's energy coefficient to be calculated. Drift velocities are measured using a magnetic deflection method. On the basis of the kinetic theory of gases these measurements are used to calculate values for the mean free path L of the electrons at unit pressure, the mean proportion η of the energy lost by an electron in a collision with a deuterium molecule, and the collisional cross section A of the molecules in collisions with the electrons. The values obtained are compared with those of Crompton and Sutton (1952) for hydrogen.

I. INTRODUCTION

Measurements have been made by Crompton and Sutton (1952) of Townsend's energy coefficient k_T in hydrogen as a function of the ratio Z/p , where p is the pressure of the gas in which the electrons are drifting under the action of the constant and uniform electric field Z . Nielsen and Bradbury (1936), using an electrical shutter method, have measured the drift velocity W of slow electrons in hydrogen. If k_T and W are known it is possible to derive, on the basis of the kinetic theory of gases, formulae which enable calculations to be made of the mean free path L of the electrons at unit pressure, the mean proportion η of the energy lost by an electron in a collision with a gas molecule, and the collisional cross section A of the molecules for collisions with the electrons. It was thought that it would be of interest to have available experimental measurements of these quantities in deuterium. In both hydrogen and deuterium the measurements were made at 15 °C.

The electron temperatures were measured by the method described by Crompton and Sutton (1952), and the drift velocities by use of the deflection of the electron stream by a transverse magnetic field. The magnetic deflection method has the disadvantage that it does not enable an absolute determination of W to be made, the value obtained depending on the manner in which the agitational speeds of the electrons are distributed about their mean value. Two possible distributions are considered, that of Maxwell and that of Druyvesteyn.

In hydrogen, where measurements of drift velocities have been made both by the electrical shutter method and by the magnetic deflection method, by comparison of the two sets of results it is possible to say that Druyvesteyn's

* Department of Physics, University of Adelaide.

distribution is that which is appropriate over a large range of values of Z/p . No previous measurements are known to have been made of drift velocities in deuterium, but it may be assumed that, for values of Z/p greater than 0.5, the value of W corresponding to the distribution of Druyvesteyn is more nearly correct.

II. APPARATUS

The diffusion apparatus used was the same as that described by Crompton and Sutton (1952). The deuterium was prepared by the method described by de Troyer, van Itterbeek, and van den Berg (1950). Heavy water was dropped slowly on to molten sodium in an evacuated vessel with the resultant evolution of deuterium. The heavy water used was stated to be 99.75 per cent. pure, and the deuterium evolved was left standing over liquid air for several hours before use in order to ensure the complete removal of all traces of heavy water vapour. In order to check the purity of the gas produced by this method, a sample of normal hydrogen was made using H_2O instead of D_2O . Values of k_1 were measured in the gas so obtained and compared with those measured when hydrogen was admitted through the walls of a heated palladium tube. Since the values were found to be identical, it was considered that the purity of the deuterium was adequate for experiments of this kind.

III. RESULTS

A summary of the results for deuterium is given in Tables 1 and 2. The various physical quantities tabulated were calculated using the following formulae (Huxley and Zaazou 1949; Crompton and Sutton 1952):

$$\begin{aligned}\text{Townsend's energy coefficient } k_T &= k_1 \text{ (Maxwell)} \\ &= 0.875k_1 \text{ (Druyvesteyn)}\end{aligned}$$

$$\begin{aligned}\text{Mean velocity of agitation of an electron } \bar{U} &= 1.06 \times 10^7 (k_1)^{\frac{1}{2}} \text{ cm sec}^{-1} \\ &\quad \text{(Maxwell)} \\ &= 1.02 \times 10^7 (k_1)^{\frac{1}{2}} \text{ cm sec}^{-1} \\ &\quad \text{(Druyvesteyn)}\end{aligned}$$

$$\begin{aligned}\text{Mean free path at unit pressure } L &= 7.11 \times 10^{-9} W(k_1)^{\frac{1}{2}}/(Z/p) \text{ cm (Maxwell)} \\ &= 7.38 \times 10^{-9} W(k_1)^{\frac{1}{2}}/(Z/p) \text{ cm} \\ &\quad \text{(Druyvesteyn)}\end{aligned}$$

Mean proportion of energy lost by an electron in a collision

$$\begin{aligned}\eta &= 1.79 \times 10^{-14} W^2/k_1 \text{ (Maxwell)} \\ &= 2.21 \times 10^{-14} W^2/k_1 \text{ (Druyvesteyn)}\end{aligned}$$

$$\text{Cross section of gas molecules for collision } A = \frac{1}{n_1 L},$$

where

n_1 = number of molecules/cm³ when $p=1$ mm
= 3.35×10^{16} at 15 °C.

$\frac{W}{K} = \frac{\text{Drift velocity}}{\text{Diffusion coefficient}} = 40.3 Z/k_1.$

TABLE 1
SUMMARY OF RESULTS FOR DEUTERIUM AT 15 °C ASSUMING MAXWELL'S DISTRIBUTION

Z/p	$W_M \times 10^{-5}$	$k_1 = k_{TM}$	$L_M \times 10^2$	$\bar{U}_M \times 10^{-7}$	$\eta_M \times 10^4$	$(\eta k_T)_M \times 10^3$	$A_M \times 10^{16}$	W/K ($p=1$ mm)
0.1	2.8	2.1 (7)	2.93	1.56	6.46	1.40	10.2	1.86
0.2	4.0	3.3 (6)	2.58	1.94	8.53	2.86	11.5	2.40
0.3	4.7	4.5 (3)	2.37	2.25	8.76	3.96	12.6	2.66
0.4	5.4	5.7 (5)	2.31	2.54	9.07	5.22	12.9	2.80
0.5	5.8	6.7 (9)	2.14	2.76	8.90	6.04	13.9	2.97
0.6	6.1	7.7 (0)	2.02	2.94	8.66	6.67	14.8	3.14
0.7	6.6	8.6 (4)	1.99	3.12	9.04	7.81	15.0	3.27
0.8	7.0	9.5 (0)	1.93	3.26	9.24	8.77	15.5	3.39
0.9	7.5	10.3	1.90	3.39	9.76	10.0	15.7	3.52
1.0	7.9	11.0	1.85	3.52	10.2	11.2	16.1	3.67
1.2	8.6	12.5	1.81	3.75	10.6	13.3	16.5	3.88
1.5	9.5	14.4	1.72	4.03	11.2	16.1	17.3	4.18
1.8	10.6	16.3	1.69	4.28	12.3	20.0	17.7	4.46
2.0	10.9	17.5	1.63	4.44	12.3	21.6	18.3	4.61
3.0	14.1	23.0	1.60	5.07	15.5	35.6	18.7	5.25
4.0	16.2	27.7	1.53	5.56	16.9	45.5	19.5	5.83
5.0	18.5	32.2	1.51	6.00	19.0	61.2	19.8	6.26
10.0		56.4						

TABLE 2
SUMMARY OF RESULTS FOR DEUTERIUM AT 15 °C ASSUMING DRUYVESTYEN'S DISTRIBUTION

Z/p	$W_D \times 10^{-5}$	k_{TD}	$L_D \times 10^2$	$\bar{U}_D \times 10^{-7}$	$\eta_D \times 10^4$	$(\eta k_T)_D \times 10^3$	$A_D \times 10^{16}$
0.1	3.1	1.90	3.36	1.50	9.82	1.86	8.89
0.2	4.4	2.94	2.96	1.87	12.9	3.79	10.1
0.3	5.2	3.97	2.72	2.16	13.2	5.24	11.0
0.4	6.0	5.04	2.65	2.45	13.8	6.96	11.3
0.5	6.4	5.95	2.46	2.66	13.4	7.97	12.2
0.6	6.8	6.75	2.32	2.83	13.2	8.91	12.9
0.7	7.4	7.56	2.29	3.00	13.8	10.4	13.0
0.8	7.8	8.32	2.22	3.14	14.2	11.8	13.5
0.9	8.3	9.02	2.18	3.26	14.8	13.3	13.7
1.0	8.8	9.64	2.15	3.38	15.5	15.0	13.9
1.2	9.5	10.9	2.07	3.61	16.1	17.6	14.4
1.5	10.5	12.6	1.97	3.88	16.9	21.3	15.2
1.8	11.7	14.3	1.94	4.13	18.6	26.6	15.4
2.0	12.1	15.3	1.87	4.27	18.6	28.5	15.9
3.0	15.6	20.2	1.84	4.88	23.4	47.4	16.2
4.0	17.9	24.3	1.76	5.37	25.6	62.2	16.9
5.0	20.5	28.2	1.73	5.77	28.8	81.3	17.2
10.0		49.4					

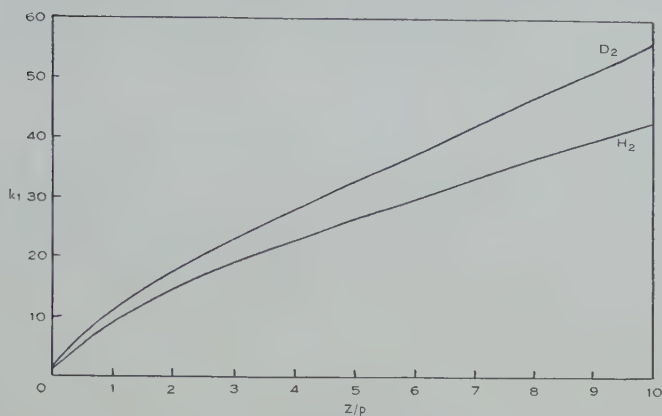


Fig. 1.—Values of k_1 in deuterium compared with those in hydrogen.

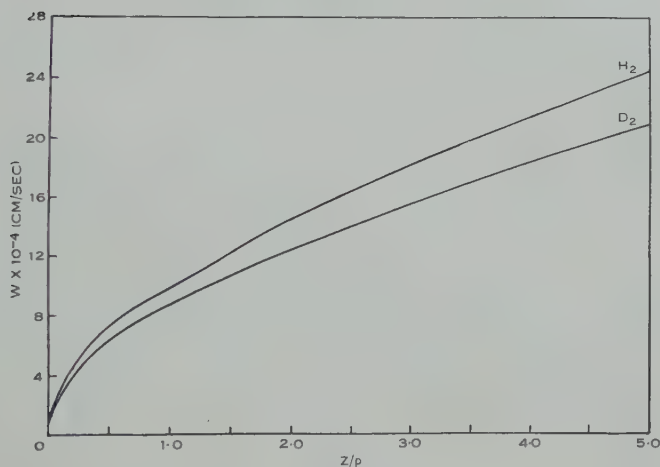


Fig. 2.—Values of W_D in deuterium compared with the values of W obtained by Nielsen and Bradbury in hydrogen.

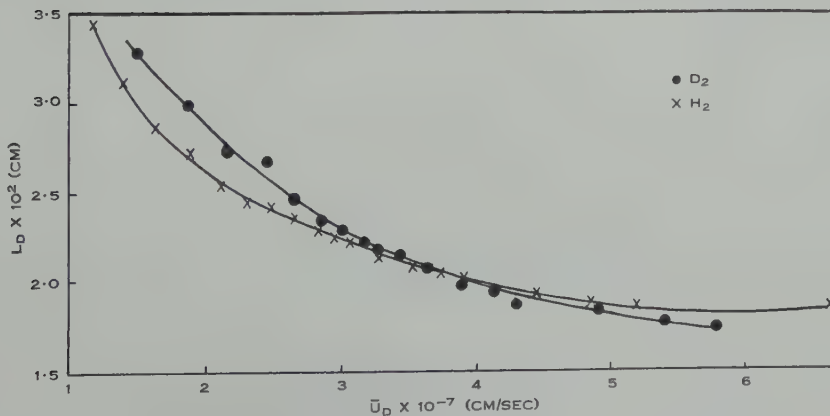


Fig. 3.—Mean free paths.

The results are shown also in the form of graphs (Figs. 1-5), together with the results of Crompton and Sutton for hydrogen, in order to facilitate comparison between the values for the two gases.

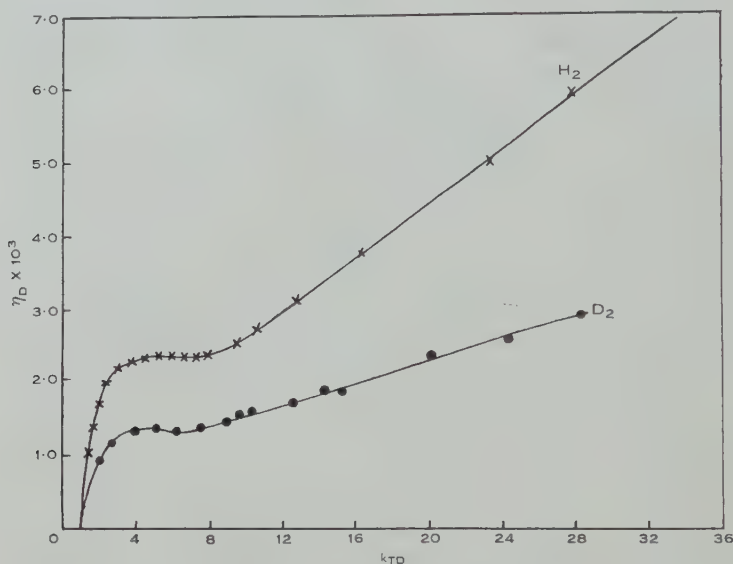


Fig. 4.—The mean proportion η of energy lost by an electron in a collision.

It can be seen that the mean free paths and hence the cross-sectional areas of the molecules of deuterium and hydrogen are almost identical for the same value of \bar{U} , which indicates that, as one would expect, these quantities are determined by the outer electronic structure of the molecule and by the nuclear

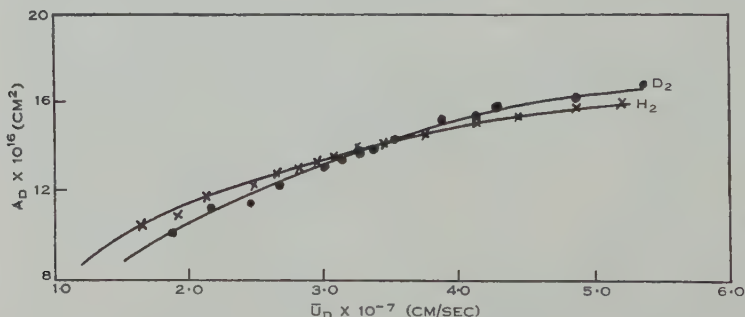


Fig. 5.—The collisional cross section A of deuterium molecules.

charges rather than by the masses of the individual nuclei. The mean proportion of the energy lost per collision is less for electrons in deuterium than for those in hydrogen.

IV. ACKNOWLEDGMENTS

The author wishes to express her thanks to Professor Huxley for encouraging these researches and to Dr. R. W. Crompton for his advice and assistance.

She would also like to thank the Commonwealth Scientific and Industrial Research Organization for the award of a studentship which enabled her to carry out this work.

V. REFERENCES

- CROMPTON, R. W., and SUTTON, D. J. (1952).—*Proc. Roy. Soc. A* **215** : 467.
HUXLEY, L. G. H., and ZAAZOU, A. A. (1949).—*Proc. Roy. Soc. A* **196** : 402.
NIELSEN, R. A., and BRADBURY, N. (1936).—*Phys. Rev.* **49** : 388.
DE TROYER, A., VAN ITTERBEEK, A., and VAN DEN BERG, G. J. (1950).—*Physica, 's Grav.* **16** : 669.

THE DISTRIBUTION OF RADIO BRIGHTNESS OVER THE SOLAR DISK AT A WAVELENGTH OF 21 CENTIMETRES

III. THE QUIET SUN—TWO-DIMENSIONAL OBSERVATIONS

By W. N. CHRISTIANSEN* and J. A. WARBURTON*

[*Manuscript received July 25, 1955*]

Summary

A distribution of solar radio brightness at a wavelength of 21 cm has been derived from observations made during the period of low sunspot activity from 1952 to 1954. The observations were made using two multiple interferometers arranged at right angles; this enabled the solar disk to be scanned in many different directions. The derived one-dimensional profiles of the quiet Sun for these various scanning directions were combined and a Fourier method adopted to derive a two-dimensional brightness distribution. The distribution shows marked limb-brightening in the equatorial zones but none in the polar regions. The contours of brightness are in general conformity with those expected from a solar atmosphere having a coronal electron density distribution of the kind proposed by van de Hulst for the period of minimum sunspot activity.

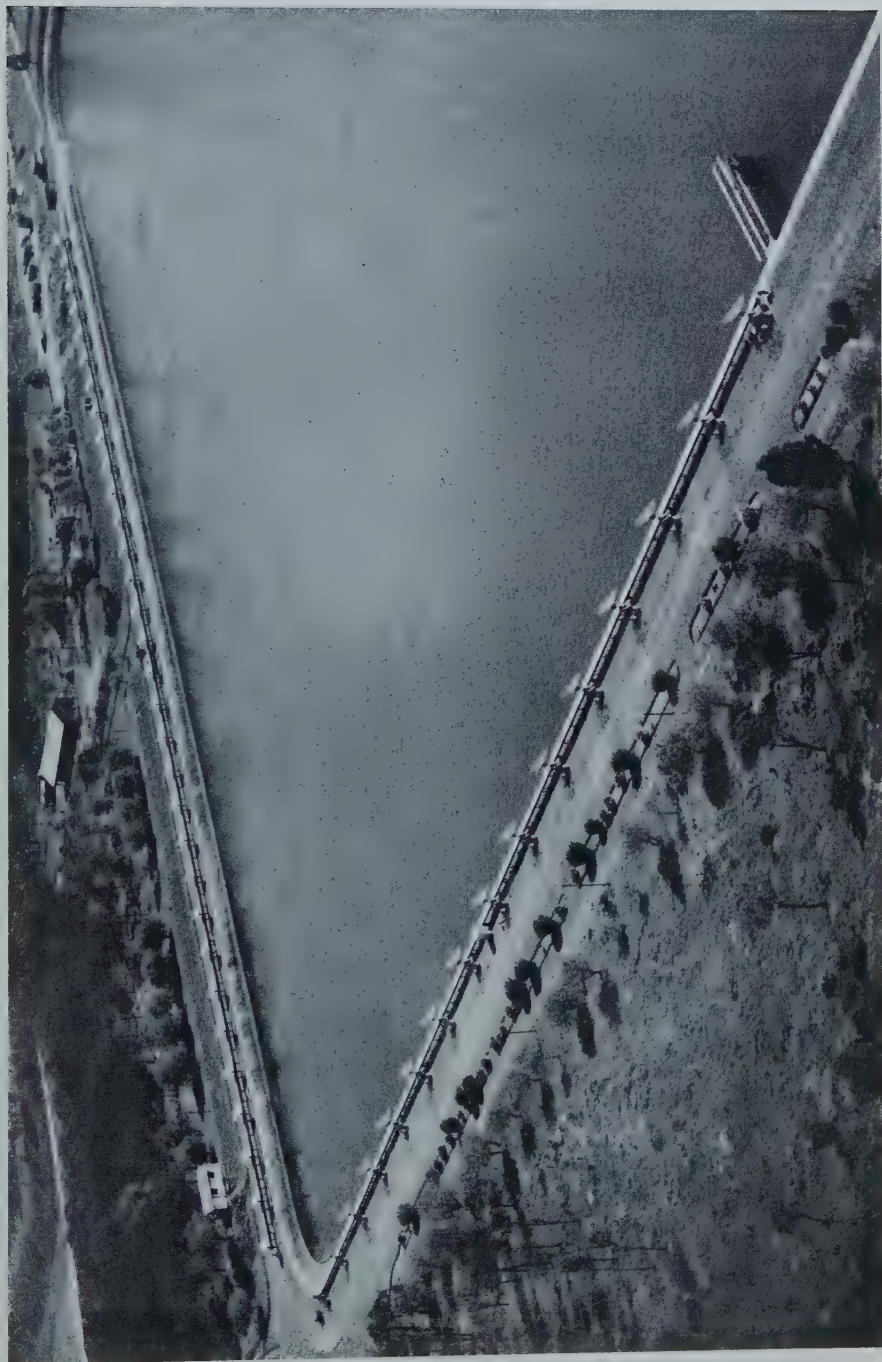
I. INTRODUCTION

It has been established in the past few years that solar radiation at decimetre wavelengths has an essentially steady background component. This "quiet" solar component has been the subject of much study recently in efforts to determine distributions of brightness across the solar disk, and work has been carried out at many radio wavelengths in the range from 6 cm to about 8 m. One of these studies was made at a wavelength of 21 cm using an aerial system which has been described in Part I of this series (Christiansen and Warburton 1953*a*). However, an essential limitation of this fixed aerial system was that the solar disk was scanned in a more or less fixed direction in a strip-wise manner by the aerial beam; consequently it was necessary to assume that the Sun had certain symmetry in order to solve the brightness distribution problem. This work, which has been described as a one-dimensional study, forms the subject matter of Part II of the series (Christiansen and Warburton 1953*b*). The presence of limb-brightening was an outstanding feature of the brightness distribution published.

In the absence of contrary information, the simplifying assumption was made that the solar disk shows circular symmetry when observed at a wavelength of 21 cm. Visual observations of the solar corona during eclipses at times of sunspot minimum have suggested that this assumption is probably not correct. In addition, radio observations during solar eclipses by Denisse, Blum, and Steinberg (1952) have given evidence of departure from circular symmetry

* Division of Radiophysics, C.S.I.R.O., University Grounds, Sydney.

RADIO BRIGHTNESS OVER THE SOLAR DISK. III



Aerial photograph of the two multiple interferometers.

of the solar disk at a wavelength of 1.78 m. The early observations at a wavelength of 21 cm by the present authors had also shown that the apparent width of the Sun seemed to change as the scanning direction varied with respect to the direction of the Sun's axis.

With the previous equipment it was not possible to scan the Sun in directions that differed greatly from the direction along the solar equator. A new multiple interferometer was built, therefore, in a direction at right angles to the earlier one. By use of the two interferometers it was possible to scan the Sun over a great variety of angles during the course of a day, and at practically any desired angle over the course of several months. From these one-dimensional scans it has been possible to reconstruct a two-dimensional brightness distribution (smoothed by the aerial) by a method of Fourier analysis.

A preliminary report of this work was given at the XIth General Assembly of U.R.S.I. where, at the same time, complementary results at a wavelength of 60 cm were reported from Cambridge. These two-dimensional results, together with those already published by O'Brien (1953) at 1.4 m wavelength, all show that the source is not circular. They also exhibit distinct differences in the degree of limb-brightening, particularly in the equatorial regions of the Sun; the maximum degree of brightening occurs at the shortest wavelength, 21 cm. The most outstanding feature of the two-dimensional distribution at 21 cm is the restricted zones in which limb-brightening is observed. No such brightening occurs in the polar regions of the sun; this is confined to the equatorial zone.

II. EQUIPMENT

The complete equipment used consisted of a grating-type interferometer of 32 equally spaced elements (described in Part I of this series) aligned approximately east-west, and a similar 16-element system arranged in an approximately north-south direction (see Plate 1). Each system has its own receiver, and all the elements of each system are connected to the receiver through equal lengths of cable. If d is the distance between adjacent elements and ψ the angle between the normal to the line of the aerials and the direction of a ray arriving at the aerial, then the angles of maximum response are given by $\sin \psi = n\lambda/d$, where n is an integer and λ the wavelength. The locus of a point making a constant angle ψ with the normal to the aerial is a small circle on the celestial sphere with the line of the aerials as axis (see Fig. 1). These small circles give the positions on the celestial sphere at which the aerial has a maximum response. If now we consider the daily motion of the Sun through such a system of beams we see (Fig. 2)* that the direction in which the Sun is scanned by the strip beams varies throughout the day. The observations on the east-west aerial were confined to a period around the hour angle (H.A.) 348° , or 12° east. Near this time the inclination of the aerial "beams" to the Sun's central meridian remains fairly constant; hence an almost exact repetition of the

* Figure 2 shows a section of a sphere on which two systems of small circles have been drawn representing the positions of the aerial beams. This sphere was used as an aid in quick analysis of the records.

pattern is produced as successive beams scan the disk. A typical record shows this (Fig. 3 (a)). With the beams produced by the north-south system, however, the observations were made from about H.A. 320° (i.e. 40° east) to H.A. 30° .

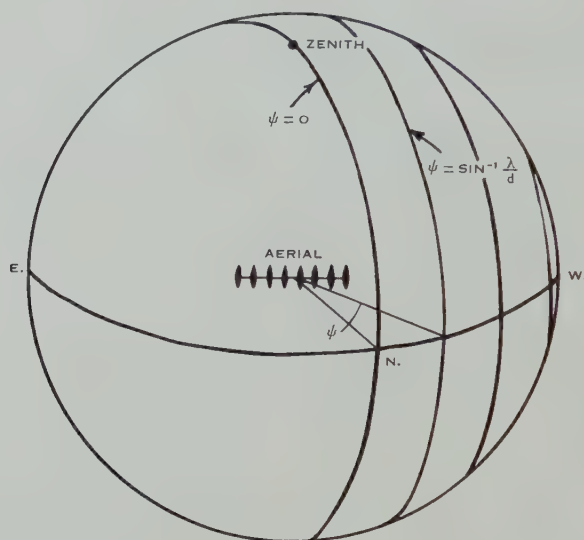


Fig. 1.—Loci of points of maximum response of a multiple interferometer as used in this work.

During this period the scanning angle changes through a range of about 50° . As the angle changes, so does the rate at which the solar disk passes through the beam. This feature is depicted clearly in the reproduced record (Fig. 3 (b)).

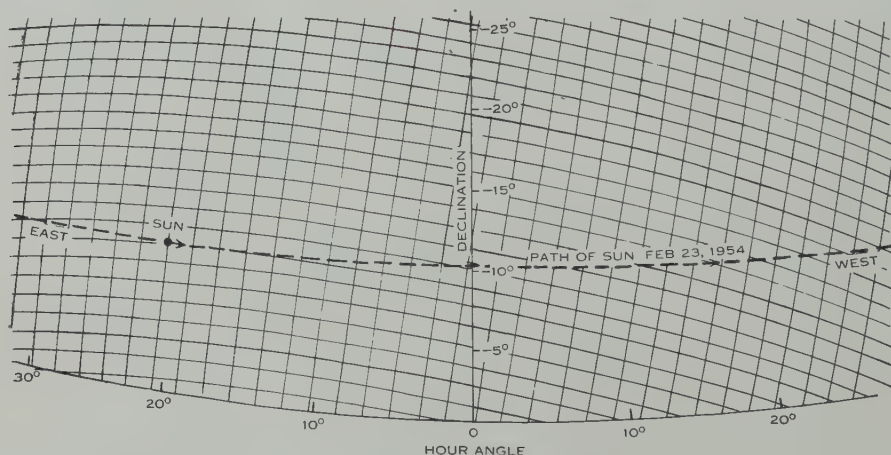


Fig. 2.—Path of the Sun through the two systems of aerial beams on one day.

Each of the aerial systems produces directly one-dimensional power *v.* time diagrams. From a knowledge of the Sun's declination the time coordinate is readily converted to the angular separation θ between the aerial beam (scanning

strip) and the centre of the optical disk. The overall length of the east-west aerial is 1028 wavelengths producing beams around midday of 3 min of arc width, while the north-south aerial of 760 wavelengths produces beams, for small values of ψ , of about 4 min of arc width. The beams actually used in the latter system depend on the Sun's declination, and for the period of observations their width ranged from 4 to nearly 5 min of arc.

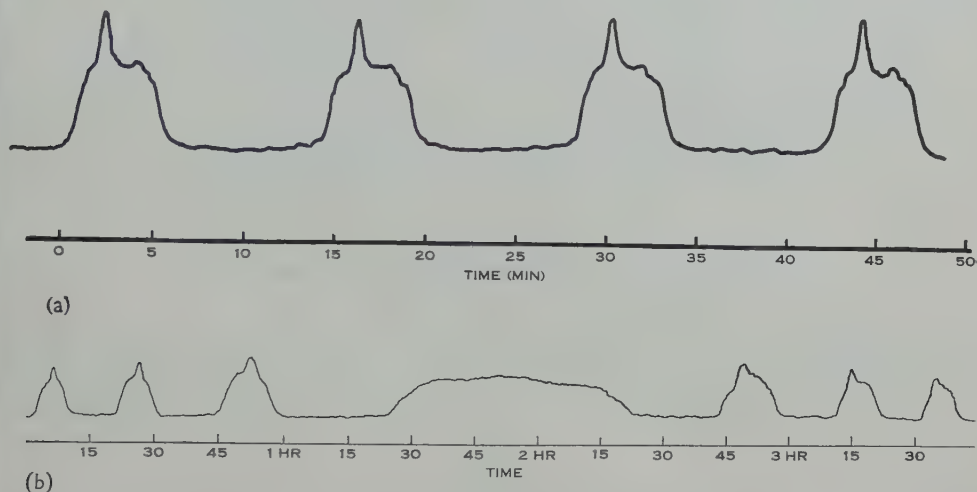


Fig. 3.—Typical daily records of the passage of the Sun through several aerial beams, (a) using the east-west interferometer, (b) using the north-south interferometer.

III. OBSERVATIONS

A series of observations has been made using both aerial systems over the period September 1953 to April 1954. We have also made observations from April 1952 to September 1953 using the east-west aerial system alone. The combination of these observations over such a long period to solve the distribution is considered to be valid because no changes in the shape or extent of the quiet Sun were detected during the period 1952–1954, for the scanning directions ϕ lying between 55° and 90° . Figure 4 shows two quiet-Sun curves observed at the same scanning angle in 1952 and 1954. Any differences are within the experimental errors.

During the observations with the two independent systems it has been possible on any one day to scan the Sun in many different directions. A typical result for October 15, 1953 is reproduced in Figure 5, where the effect of scanning the Sun in two directions differing by 68° is shown. The departure from circular symmetry of the quiet Sun is clearly illustrated by these observations.

By making use of the seasonal variation of the Sun's orientation with respect to the two sets of aerial beams, it has been possible to obtain records covering 140° out of 180° range of scanning angle, as indicated in Figure 6. (The scanning angle ϕ is defined as the angle between the Sun's central meridian and the normal to the aerial beam.)

The geometry of the two systems is such that the range of angles covered by one system is complementary to that of the other, the two ranges approaching to within 6° of one another. Obviously, the time taken for the Sun to drift

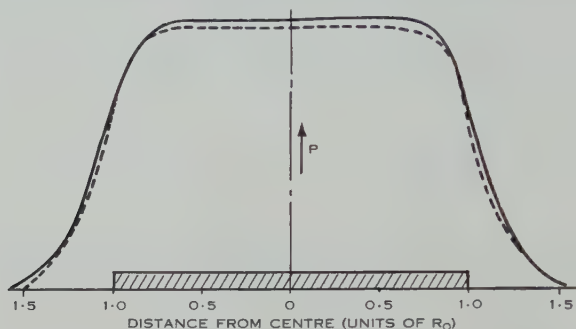


Fig. 4.—One-dimensional quiet Sun profiles for the same scanning direction $\phi = 58^\circ$. ---- 1952, — 1954. P is the power received in arbitrary units.

through a beam depends on its declination, on the angle that the beam makes with the Sun's diurnal path, and on the Sun's semi-diameter. It is therefore

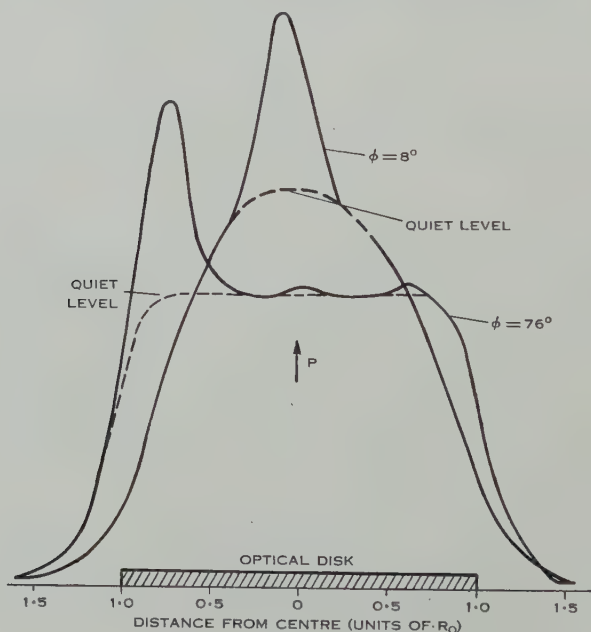


Fig. 5.—One-dimensional profiles showing the result of scanning the Sun in different directions on the same day.

necessary to normalize the time scale of each record. The next step is to subtract contributions due to localized bright regions. For this purpose we superposed several diagrams, taken on different days for each scanning direction, and drew

the lower envelope (see Part II). Figure 7 shows the result of superimposing several days' observations for an average scanning angle of 28° . The deduction

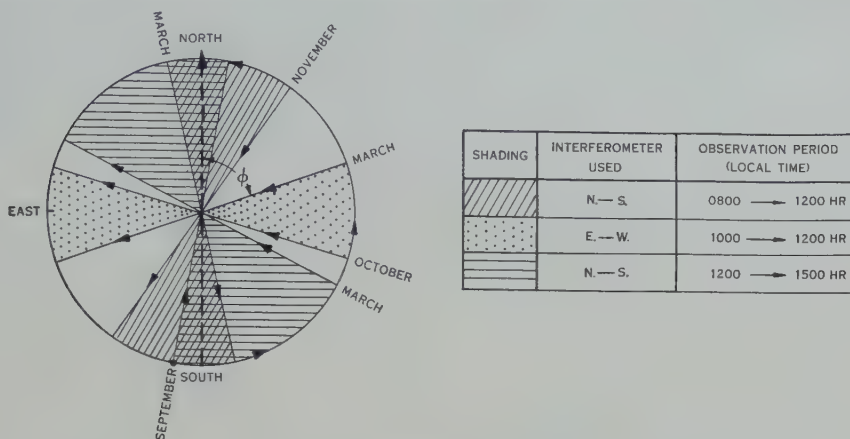


Fig. 6.—Range of scanning angles ϕ covered by the two aerial systems.

of the quiet Sun level from such superimposed diagrams was usually not very difficult because during the period of observation localized regions of enhanced brightness on the Sun were not frequent.

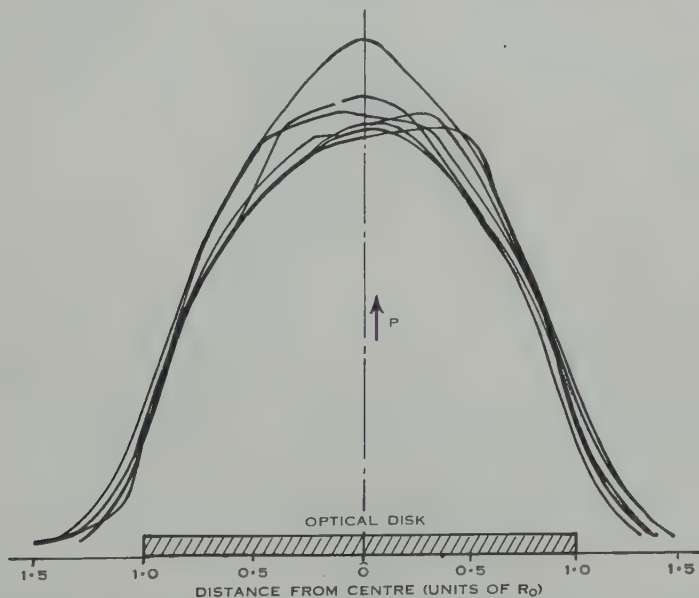


Fig. 7.—A succession of daily records for scanning directions near $\phi = 28^\circ$ superimposed to show the contribution of the quiet Sun.

Now, each of the one-dimensional curves is symmetrical; also the same result is obtained for each direction which makes the same given angle with the Sun's central meridian. Such observations indicate that the true brightness

distribution should possess quadrant symmetry. For this reason we have restricted the analysis to one quadrant of the distribution, $\varphi=0^\circ$ to $\varphi=90^\circ$. Figure 8 shows the set of observed one-dimensional scans used in the analysis. These are for angles of $\varphi=0, 28, 42, 51, 57, 69$, and 82° , and each is the average curve over a range of angles of $\varphi \pm 6^\circ$. These scans represent the one-dimensional or line integrated brightness distributions over the solar disk, smoothed by the effect of the aerial beam.

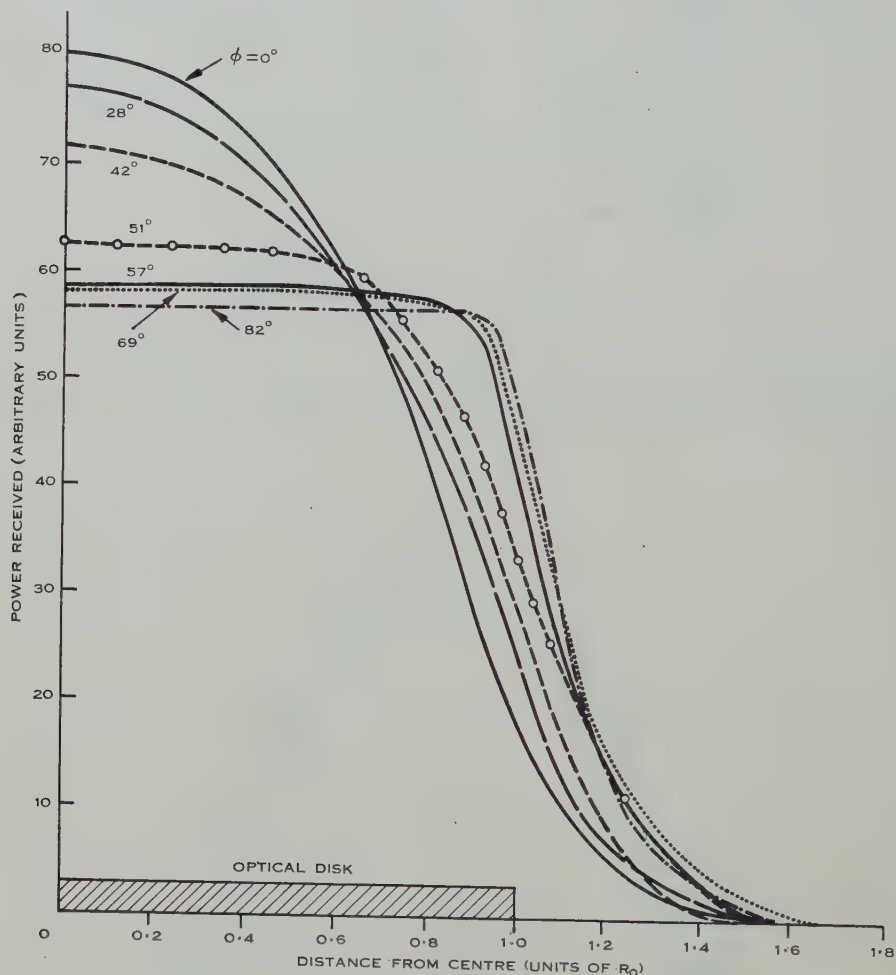


Fig. 8.—One-dimensional quiet Sun profiles used in the analysis for several scanning angles φ .

IV. THE DERIVATION OF THE BRIGHTNESS DISTRIBUTION FROM THE OBSERVATIONS BY A METHOD OF FOURIER ANALYSIS

The way in which a two-dimensional radio brightness distribution may be derived from a number of one-dimensional scans is not obvious. However, rather similar two-dimensional problems have arisen in the field of crystallography, and solutions for these problems, using the methods of Fourier analysis,

have been found. Such a method of solution has been applied to the problem of a radio brightness distribution by O'Brien (1953). It is well known that the two-aerial interferometer, as used by O'Brien and others, measures the Fourier components of the strip integration of the real distribution, and that the frequency of the component being observed is determined by the aerial spacing. With a grating-type interferometer, as used in this work, the one-dimensional distribution is observed directly. The Fourier component of highest frequency contained in the distribution is of course determined by the length of the aerial system. Hence, if we take the Fourier transform of this one-dimensional distribution, the available information is then in a similar form to that of O'Brien and the subsequent solution of the two problems becomes identical. A full discussion of this problem is in preparation by Smerd and Wild.

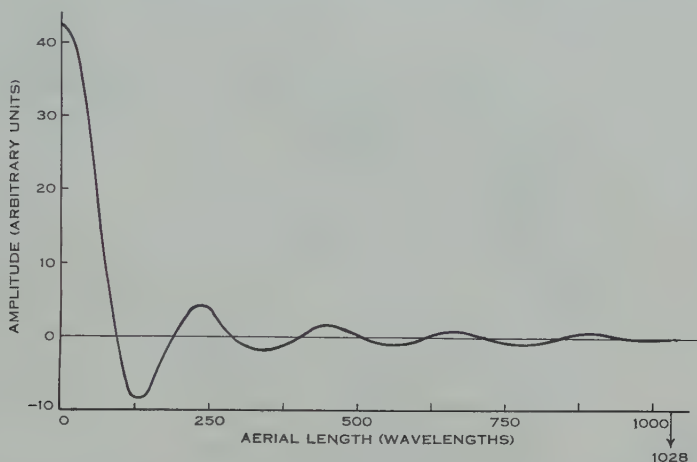


Fig. 9.—Cosine Fourier transform of the one-dimensional profile of the quiet Sun in the direction $\phi=75^\circ$.

The procedure adopted in the derivation of the two-dimensional distribution was as follows. The cosine Fourier transform of each one-dimensional distribution was first determined numerically. Only the even (cosine) terms of the Fourier components are present because each of the observed distributions is symmetrical. A typical cosine transform of one of the observed distributions is shown in Figure 9, and it is seen that both positive and negative values appear. Each of these transforms has a finite high frequency cut-off which is inversely proportional to the width of the aerial beam used in the observation. Since it seemed undesirable in the analysis to have higher resolution in some directions than in others (the beam widths varied from less than 3 min of arc to slightly more than 4 min of arc), some transforms have been degraded so that all correspond to the use of a uniformly fed aperture aerial with a half-power beam width of 4.3 min of arc. By plotting each of these "rippled" transforms radially in the direction corresponding to its scanning direction, a two-dimensional Fourier transform diagram was constructed. Contours of equal amplitude were drawn. The amplitude at the centre of the diagram, being the zero-frequency

component, is of course given by the area under the one-dimensional distribution curves and is the same for each curve. We now strip integrate this diagram, the strip summations being perpendicular to the scanning directions. The cosine Fourier transforms of these strip integrals give the radial cross sections of the brightness distribution in their corresponding scanning directions. The two-dimensional distribution is then obtained directly by plotting these radial cross sections and drawing contours of equal brightness. Figure 10 shows the derived

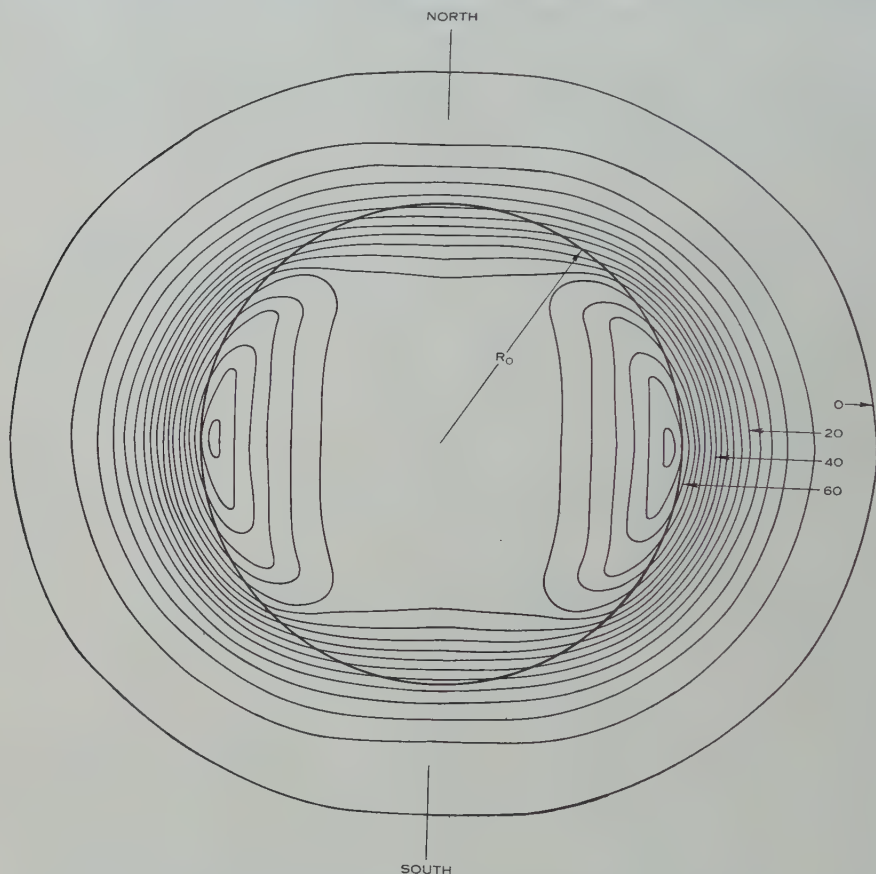


Fig. 10.—Derived two-dimensional radio brightness distribution. The contours are at equal intervals of 4×10^3 °K. The central brightness temperature is 4.7×10^4 °K, and the maximum peak brightness is 6.8×10^4 °K.

distribution. In the absence of errors in the calculations, line integration of the distribution in Figure 10 should reproduce the observed one-dimensional curves in Figure 8. In fact this process produces curves which differ from those in Figure 8 by less than 3 per cent. at any point.

If we know the total flux received from the quiet Sun, it is possible to assign absolute magnitudes of brightness temperature to the contours of Figure 10. For this, the apparent disk temperature (T_d) of the quiet Sun was taken as 7.0×10^4 °K. This temperature was deduced from the lowest limits of solar

RADIO BRIGHTNESS OVER THE SOLAR DISK. III



Photograph of the Sun at the total eclipse June 30, 1954 (M. Waldmeier).

radio frequency flux as published in the I.A.U. Quarterly Bulletin of Solar Activity. During 1953, the lowest limits of flux coincided with the complete absence of local bright regions on the solar disk as seen on our records, and hence these values may be taken to refer to the quiet Sun.

V. THE DISTRIBUTION : FEATURES

The most outstanding feature of the distribution is the departure from circular symmetry. This characteristic has been found elsewhere at lower frequencies. Eclipse observations by Denisse, Blum, and Steinberg (1952) at 169 Mc/s were interpreted to demonstrate marked departure from a circularly symmetrical source. The source was shown to extend 1.6 times further in the equatorial direction than in the direction of the solar axis. More recently O'Brien (1953), from observations at 1.4 m, has given a two-dimensional distribution which extends 1.2 times further along the equatorial direction than in the direction of the solar axis. The lower contours of the distribution in Figure 10 also show this feature. If we choose the contour at which the brightness temperature has fallen to half the centre value, then the radial extent is $1.25 R_0$ in the equatorial direction, whereas it is $0.94 R_0$ in the direction of the poles.

The contours of Figure 10 are at equal intervals of brightness temperature $T_b = 4.0 \times 10^3$ °K, the brightness temperature at the centre being 4.7×10^4 °K.

The limb-brightening reported in Part I is found to be confined to the equatorial regions. The peak of brightness appears to fall just inside the optical limb, rather than just beyond, as predicted theoretically. This discrepancy may not be significant, however, as the effect of the aerial beam has not been removed, and it is readily seen that smoothing by the aerial beam could cause the peak to fall inside its true position. The value of the peak in the derived brightness temperature is 1.46 times the value at the centre, i.e. 6.8×10^4 °K, and falls off to 1.2 times in a direction 60° from the pole. At angles less than 30° no limb brightening occurs at all (see Fig. 11). Another effect of the aerial beam, which has been pointed out in Part II, is to make the real values for the peak brightness indeterminate. Hence the true brightness temperatures at the optical limb may be considerably higher than the values quoted.

The heliographic latitude at which limb-brightening ends is about 55° , which is approximately the latitude at which changes in the coronal structure can be seen at times of sunspot minimum. This feature is shown very clearly in the remarkable photograph (Plate 2) of the solar corona obtained by Waldmeier* during the eclipse of June 1954. At this time the sunspot activity cycle was in its minimum phase, and the Sun was exceptionally free from active regions; hence the photograph is very well suited for comparison with the derived quiet Sun distribution. The photograph shows that the structure and extent of the corona change at latitudes around 50° to 60° ; at latitudes higher than these, polar plumes appear, while near the equator the corona extends to form well-defined wings. The corresponding feature at the 55° latitude in the derived

* We are indebted to Professor M. Waldmeier for permission to reproduce this photograph.

brightness distribution (Fig. 10) is the change from limb-brightening to limb-darkening.

The greater radial extent in the equatorial direction compared with that in the direction of the poles is also evident in the photograph. This feature can be compared in more detail with the radio brightness distribution, since the photographic image shows an apparent limit to the corona, and this limit is, in fact, a brightness contour for the corona. If we compare this contour with those of Figure 10, we find that the 8000 °K contour follows closely the outline of the corona as it appears in the photograph. (The "kinks" in the optical picture would not be expected to appear in the radio picture because of the

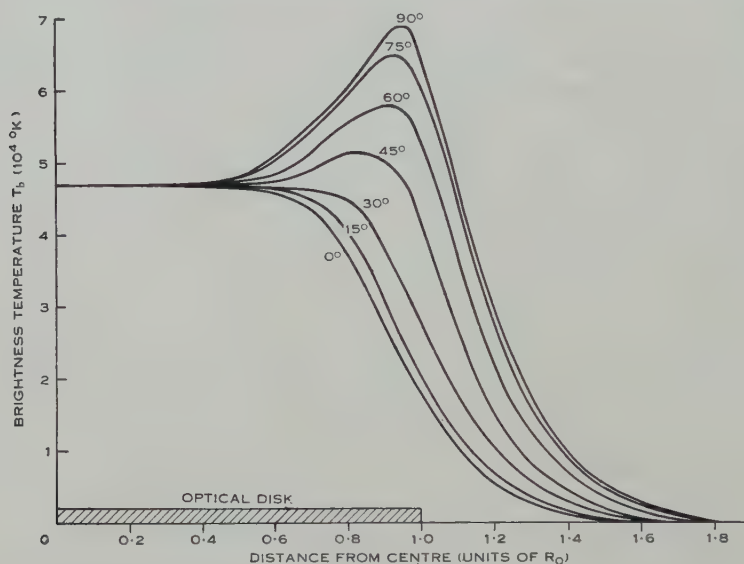


Fig. 11.—Derived radial brightness distributions in several directions with respect to the pole of the Sun.

smoothing effect of the radio aerial.) In both cases the part of the solar corona that is depicted is optically thin. This means that the contours represent lines of constant numbers of electrons integrated along the line of sight in the optical case, and (assuming uniform coronal temperature) constant means of the square of this number over regions around the line of sight in the radio case. The agreement between radio and optical contours (within the limitations imposed by the smaller resolution of the radio telescope) is satisfactory. It highlights the complementary nature of radio and optical studies of the Sun, and it may be taken as substantiating prevailing ideas on the origin of quiet Sun radiation at centimetre wavelengths.

VI. THE DISTRIBUTION : DISCUSSION

It is generally agreed that at decimetre wavelengths pronounced limb-brightening is likely to occur because of the large increase in the coronal contribution to the radiation at the limb compared with that at the centre of the disk.

The radiation in the centre emanates mainly from the chromosphere. This is because the opacity of the corona in the line of sight is relatively small. Near the limb, however, the emergent ray is confined to the corona and, because the ray path in this region is longer, the opacity is greater. Since the electron temperature of the corona is much higher than that of the chromosphere, the brightness temperature observed in these regions may be higher.

In the two-dimensional distribution presented in Figure 10 this brightening occurs in the equatorial region, the degree of brightness and the position at which it occurs conforming qualitatively with the values predicted theoretically from simple solar models.

The observed brightness distribution, by itself, cannot be used to derive the distribution of temperature and pressure anywhere in the solar atmosphere. Even if similar information at other radio frequencies is available, it appears that a determination of these quantities requires the addition of information from optical solar observations. One procedure that can be adopted is to postulate model solar atmospheres which are consistent with what is known from optical observations and to compare the theoretical brightness distributions from such sources with that derived from observation. This comparison, however, is not usually direct. It must be remembered that the observations have been made with an aerial system of some particular resolving power, and unresolved features in the emitting surface of the Sun cannot be reproduced from the observations. Near the limb of the Sun's disk, where rapid changes in brightness with radial distance may occur, smoothing by the aerial system may be expected, and observed peak values of brightness temperature may fall well below the true values. A comparison between observed and theoretical brightness distributions should therefore be made using distributions which have been smoothed to the same extent.

In this paper we shall not attempt a detailed comparison between the observed and theoretical brightness distributions; we shall simply indicate that the brightness temperatures derived from the observations, and the temperatures expected from simple models based on available optical and radio data, are not grossly at variance.

For the appropriate period, that is, of sunspot minimum, the only values available for the electron density in the Sun's corona are those of van de Hulst (1949), who has tabulated the densities for both the equatorial and polar regions of the corona. If we employ these values in a simple solar model in which a uniform temperature corona is assumed, and consider two points at a radius of 1.05 times the optical radius of the Sun where only coronal emission will be observed, we find that, for a 10^6 °K corona, the brightness temperature at the equator would be 5.7×10^4 °K and at the pole, 2.0×10^4 °K.* These values are in order of magnitude agreement with those of 4.7×10^4 °K and 1.4×10^4 °K derived from our observations.

The derived brightness temperature for the central ray is of importance because in this region one may expect that the finite beam width of the aerial

* The authors wish to thank Mr. S. F. Smerd for some of his unpublished computations which were used in these calculations.

will produce little error in the measurement. If we use the same model corona as before, we find that about 15 per cent. of the radiation in the central ray originates in the corona. Hence the chromospheric component has an effective brightness temperature of 4×10^4 °K. To relate this value with electron temperatures and pressures in the chromosphere one may combine the result with those at other frequencies and also with the results of optical observations, as was done, for example, by Piddington (1954).

(a) *Long-term Variations in the Quiet Sun*

The radio-frequency brightness distribution in Figure 10 has been obtained from observations made at or near to a minimum of the sunspot cycle. As the region of origin of radiation at this wavelength lies partly in the corona, it is possible that the brightness distribution may change with the solar cycle as visual observations of the corona strongly suggest. Evidence of changes in the radio emission from the quiet Sun during the solar cycle was found by Covington (1949) and Christiansen and Hindman (1951). However, the latter results have been challenged by Piddington and Davies (1953), who consider that previous estimates of the quiet Sun radiation, made during sunspot maximum, actually involved a considerable proportion of radiation from old active regions on the Sun. They conclude that there has been no significant change in the output of the quiet Sun. Further measurements of the radio brightness distribution at other times in the solar cycle are obviously needed to resolve the problem of whether the quiet Sun does or does not change during the sunspot cycle.

VII. CONCLUSIONS

The measurements of brightness distributions at a wavelength of 21 cm have shown that at sunspot minimum the brightness shows differences between equatorial and polar regions which agree qualitatively with the eclipse photographs of the Sun at sunspot minimum. Further measurements will be required to determine whether there are significant changes in the brightness distribution during the course of a solar cycle. From visual observations this would be expected, and we must wait until the cycle moves away from its minimum phase before we can verify the changes which have been predicted.

VIII. ACKNOWLEDGMENTS

The authors wish to acknowledge the help of Mr. G. Swarup in the computations involved in this work, and of Mr. G. Chandler and Mr. C. Fryar in the taking of the observations. We wish to thank also Mr. J. P. Wild and Mr. S. F. Smerd for their helpful criticism of the paper.

IX. REFERENCES

- CHRISTIANSEN, W. N., and HINDMAN, J. V. (1951).—*Nature* **167**: 635.
 CHRISTIANSEN, W. N., and WARBURTON, J. A. (1953a).—*Aust. J. Phys.* **6**: 190.
 CHRISTIANSEN, W. N., and WARBURTON, J. A. (1953b).—*Aust. J. Phys.* **6**: 262.
 COVINGTON, A. E. (1949).—*Proc. Inst. Radio Engrs.*, N.Y. **37**: 407.
 DENISSE, J. F., BLUM, E. J., and STEINBERG, J. L. (1952).—*Nature* **170**: 191.
 VAN DE HULST, H. C. (1949).—*Nature* **163**: 24.
 O'BRIEN, P. A. (1953).—*Mon. Not. R. Astr. Soc.* **113**: 597.
 PIDDINGTON, J. H. (1954).—*Astrophys. J.* **119**: 531.
 PIDDINGTON, J. H., and DAVIES, R. D. (1953).—*Nature* **171**: 692.

SOLAR BRIGHTNESS DISTRIBUTION AT A WAVELENGTH OF 60 CENTIMETRES

I. THE QUIET SUN

By G. SWARUP and R. PARTHASARATHY*

[*Manuscript received July 15, 1955*]

Summary

A multiple-element interferometer has been employed to determine one-dimensional distributions of radio brightness over the quiet Sun at a wavelength of 60 cm for scanning directions varying from 90° to 60° with respect to the central meridian of the Sun. These observations have been compared with measurements by other workers at the same, or nearly the same, wavelength. The present observations are reasonably consistent with the two-dimensional brightness distribution derived recently by O'Brien and Tandberg-Hanssen with a two-aerial interferometer, but do not agree with the earlier results of Stanier at the same wavelength. The disagreement, largely the absence of the theoretically predicted limb-brightening in Stanier's results, may reflect actual changes in the Sun over the solar cycle. However, the possibility of localized disturbed regions affecting Stanier's results for the quiet Sun cannot be eliminated.

I. INTRODUCTION

Solar radiation at decimetre wavelengths is known to consist of a steady background, attributed to the quiet Sun, and variable components that are emitted by localized disturbed regions on the Sun.

The brightness distribution of the quiet Sun is of considerable interest because of the information it may give on conditions in the solar chromosphere and corona. On the basis of model solar atmospheres derived from optical observations several workers have suggested that the Sun should show limb-brightening at decimetre wavelengths.

The first attempts to investigate this were made using eclipse observations. Christiansen, Yabsley, and Mills (1949) made detailed observations of an eclipse in 1948 to investigate the radio brightness distribution over the solar disk at a wavelength of 50 cm. But their results were not conclusive as it was found difficult to isolate completely the "quiet" component of solar radiation. In addition, the eclipse curve was insensitive to widely different, assumed distributions. Stanier (1950) used a two-aerial interferometer to derive a distribution across the quiet Sun at a wavelength of 60 cm. This showed no limb-brightening. On the other hand, limb-brightening at shorter wavelengths has been deduced from the recent observations at 3 cm by Alon, Arsac, and Steinberg (1953,

* Division of Radiophysics, C.S.I.R.O., University Grounds, Sydney.

1955), at 10 cm by Covington and Broten (1954), and at 21 cm by Christiansen and Warburton (1953*b* ; 1955).

It has been found difficult to explain Stanier's distribution on the basis of currently accepted solar models. With the technique employed by him, however, there was no direct way of eliminating with certainty the slowly varying component of solar radiation. The observations were taken near the maximum phase of the solar cycle and it was likely that a number of localized, disturbed regions may have existed on the Sun. The possibility of unsuspected, localized radio-bright regions on the Sun affecting Stanier's result prompted us to use an independent technique to explore the brightness distribution across the quiet Sun at the same wavelength, by means of a 32-element interferometer. The usefulness of a multiple-element interferometer of high resolving power for a study of the different components of the solar radiation in the range of decimetre wavelengths has been shown by the results of Christiansen and Warburton (1953*a*, 1953*b*) at 21 cm.

The preliminary results of our observations of the quiet component of solar radiation, which showed considerable limb-brightening, were reported at the XIth General Assembly of U.R.S.I. At the same time it was reported from Cambridge that a recent experiment with a two-aerial interferometer also demonstrated limb-brightening at a wavelength of 60 cm. Brief reports of the results with these two independent techniques have been published recently (O'Brien and Tandberg-Hanssen 1955 ; Swarup and Parthasarathy 1955). The two results are compared here.

The results emerging from our observations of radiation from localized disturbed regions on the Sun will be submitted in another paper.

II. EQUIPMENT

The 32-element interferometer has been described in detail by Christiansen and Warburton (1953*a*). It consists of 32 paraboloidal aerials, each 66 in. in diameter, arranged at intervals of 23 ft in a straight line lying approximately in an east-west direction. The aerials are joined to a sensitive radio receiver by a branching system of two-wire open-transmission lines. For the present study, the aerial system was converted to operate at a wavelength of 60 cm.

The interferometer gives a number of fan-shaped beams which, at a wavelength of 60 cm, are separated by 4.9° . The calculated half-power width of the central beam in an east-west direction is 8.2 min of arc. Owing to rotation of the Earth, each fan-shaped beam scans the solar disk stripwise, producing a number of one-dimensional distributions across the solar disk every day. Some of the daily records are shown in Figure 1.

The aerial pattern was experimentally determined by observing the response of the radio source of Cygnus whose width at this frequency is much smaller than that of the aerial beam. The half-power width of the interferometer beam thus determined was found to be 8.7 min of arc, instead of the calculated value of 8.2 min of arc. The beam width was also measured by determining

the excess radiation of a strong, localized, bright region (see Fig. 1 (c)) over that of the steady background due to the quiet Sun. This also gave the half-

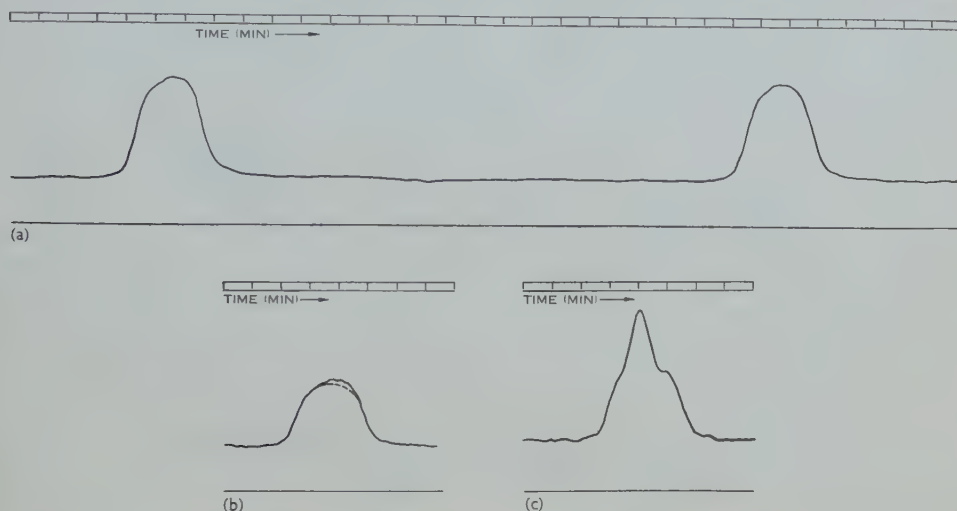


Fig. 1 (a).—A typical record of the passage of the Sun through successive aerial beams.

Fig. 1 (b).—A record for July 6, 1954 showing a small localized bright region. The dotted line indicates the quiet Sun.

Fig. 1 (c).—A record for December 12, 1954 showing a large bright region.

power width as 8.7 min of arc. The aerial pattern thus determined is shown in Figure 2, together with the theoretical curve for a beam of 8.2 min of arc

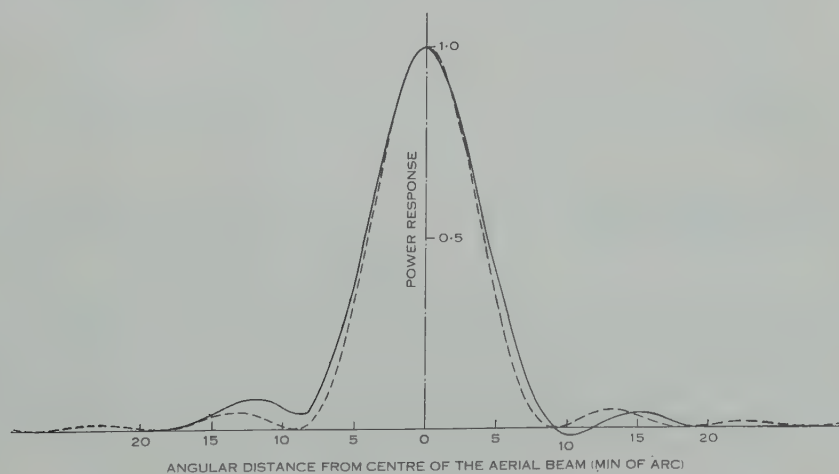


Fig. 2.—Theoretical aerial pattern (broken line) compared with the derived response of a localized bright region on the solar disk (full line).

width. The two are reasonably similar. Differences in the value of side lobes could be due to small experimental errors.

III. OBSERVATIONS

The contribution of radiation from the localized bright regions on the solar disk can be separated from the steady base-level component due to the quiet Sun by superimposing a number of daily records, using the method described by Christiansen and Warburton (1953*b*). The lower envelope of the records is taken to be the one-dimensional distribution across the quiet Sun. The daily

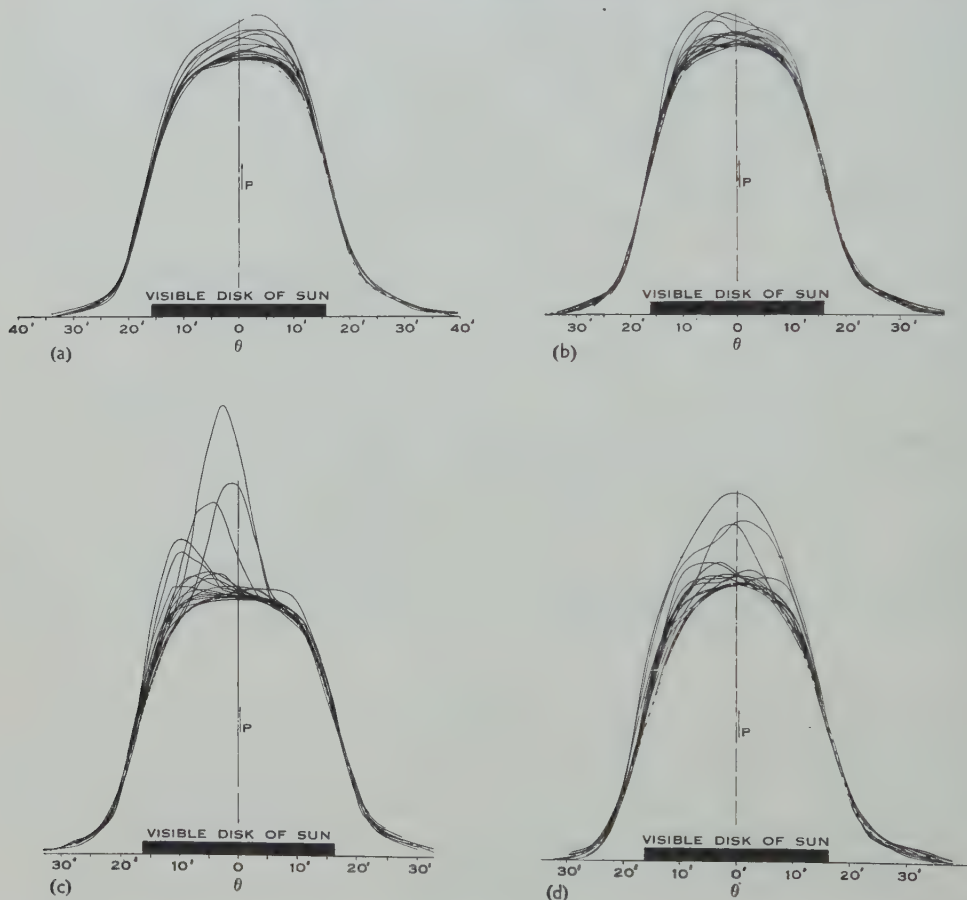


Fig. 3.—Superimposed daily records of the one-dimensional brightness distribution across the Sun for several periods during July 1954 to March 1955. (a) July 18–August 5, $\varphi=3^\circ$; (b) August 9–September 1, $\varphi=11.5^\circ$; (c) December 15–January 3, $\varphi=-3^\circ$; (d) February 7–March 4, $\varphi=26^\circ$. φ is the mean angle between the scanning direction and the central meridian of the Sun; it changed by about 6° in each observing period. P is the power received in arbitrary units. θ is the angle in minutes of arc between the centre of the visible disk and the centre of the aerial beam.

records were superimposed in groups of about 20 for several periods during July 1954 to March 1955. The superimposed records are shown in Figure 3. The estimated base level is indicated by a dotted line. On several days during this period the observed records coincided with this estimated base level. At such

times optical observations showed that there was no sign of sunspots or chromospheric faculae on the Sun. Hence we may conclude that the lower envelope of the superimposed records does, in fact, refer to the quiet Sun.

The direction of scanning of the solar disk by the aerial beam changed during the period of recording as a result of a variation in the position angle P of the solar axis. On normalizing the one-dimensional distributions across the quiet Sun, shown in Figure 3, to the same reference axes, it is noted that the diagrams become narrower in width when the scanning direction changes from 90° to 60° with respect to the central meridian of the Sun. This is demonstrated in Figure 4 and implies that at a wavelength of 60 cm the solar disk is not circular, but has a maximum width in the equatorial region. This is consistent with the findings of Christiansen and Warburton (1955) at 21 cm wavelength and of O'Brien (1953) at 1.5 m wavelength. Recent two-dimensional results of O'Brien and Tandberg-Hanssen (1955) at 60 cm confirm this.

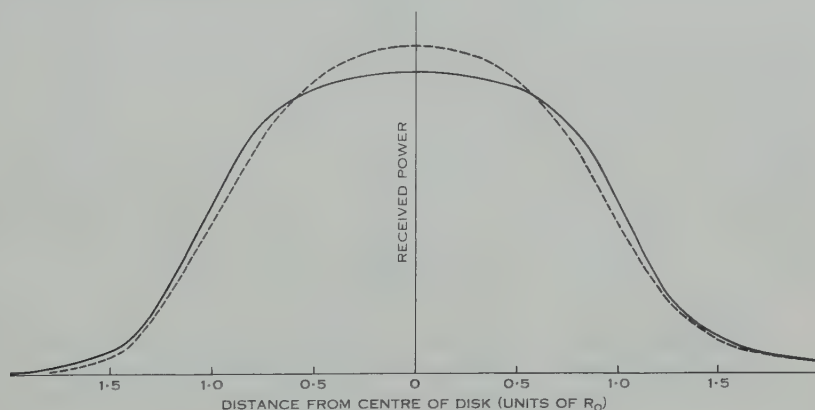


Fig. 4.—One-dimensional distributions over the quiet Sun (normalized to the same area) for scanning directions of 90° (full line) and 64° (broken line) with respect to the Sun's central meridian.

IV. COMPARISONS AND DISCUSSION

The present observations covered only a limited range of scanning directions. It was, therefore, not possible to derive a two-dimensional picture of the Sun without making assumptions about the Sun's shape. However, we can compare our one-dimensional distributions with the results derived by other workers at the same or nearly the same wavelength.

The comparisons to be made are with the results of Stanier (1950) and O'Brien and Tandberg-Hanssen (1955) and with the eclipse observations by various workers. These other measurements were made at different phases of the solar cycle, and this must be taken into account in any comparison. It may be mentioned here that in the present investigations the base level for the apparent disk temperature of the Sun at 60 cm was estimated to be 3.8×10^5 °K., whereas at the time of Stanier's measurements it was estimated to be 5.4×10^5 °K.

(a) Comparison with the Results of Stanier

Stanier used a variable-spacing two-aerial interferometer to derive the solar brightness distribution at a wavelength of 60 cm. He scanned the solar disk in the equatorial direction. To compare his results with ours, it was necessary to allow for a difference in the response for the received Fourier components in the two aerial systems. The overall spacing of Stanier's and our aerial systems is similar. That means the same range of Fourier components was measured in the two observations. However, the weighting of the Fourier components differs in the two aerial systems. The variable-spacing two-aerial interferometer attaches equal weights to all components, while the multiple-

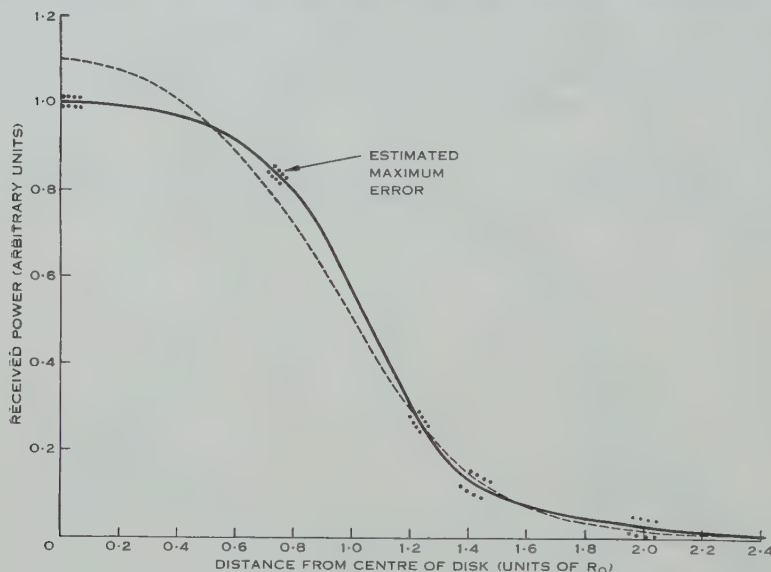


Fig. 5.—Comparison of our one-dimensional distribution over the quiet Sun for equatorial scanning direction (full line) with that derived on smoothing Stanier's distribution with our aerial beam (broken line).

element interferometer attaches weights which decrease steadily towards the cut-off frequency. Allowing for this difference, the comparison was made in two different ways.

First, the one-dimensional distribution obtained by us for the equatorial scanning direction was compared with that of Stanier's. The latter was not published in his paper. It was here derived by smoothing the published radial brightness distribution, which was derived on the assumption of circular symmetry, by our aerial pattern. This is an equivalent process to reducing the weights of the Fourier components in Stanier's distribution to those of ours. The two curves are shown in Figure 5; they differ by amounts which considerably exceed our estimated error.

The effect of this disagreement shows out more clearly when the corresponding radial brightness distributions are compared. For the purpose of this comparison we derived the radial brightness distribution from our equatorial

scan by assuming, like Stanier, a circularly symmetrical Sun. The different response of the two aerial systems was here allowed for by restoring the Fourier components of our one-dimensional distribution. This was done by an approximate method described by Bracewell (1955). Our radial brightness distribution shows considerable limb-brightening (Fig. 6) whereas Stanier's results show limb-darkening.

(b) *Comparison with the Results of O'Brien and Tandberg-Hanssen*

Recently O'Brien and Tandberg-Hanssen (1955) obtained a two-dimensional distribution across the Sun at 60 cm by a series of measurements with a variable-spacing two-aerial interferometer in which the Sun was scanned in several directions. Their distribution, which is non-circular, exhibits limb-brightening. In order to compare our one-dimensional distributions with their two-dimensional

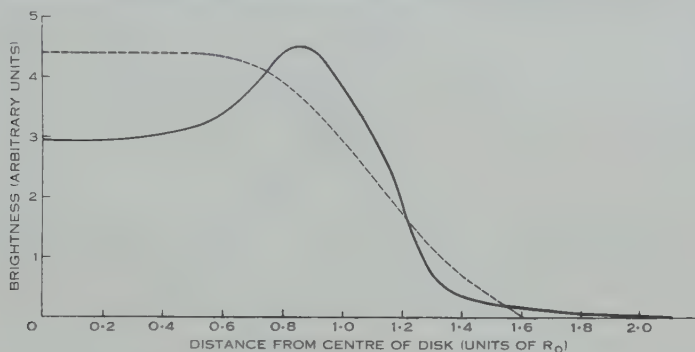


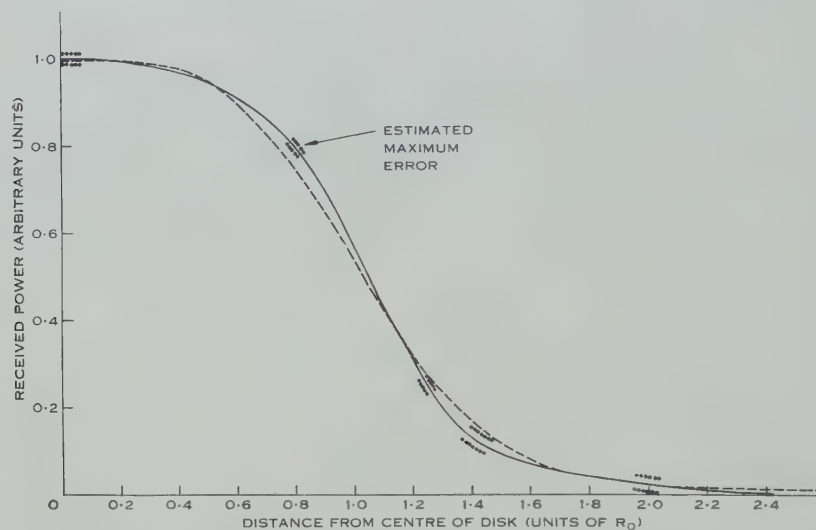
Fig. 6.—Radial brightness distribution across the solar disk at 60 cm derived on assumption of circular symmetry. The full line is the one derived by us and the broken line that by Stanier. Both are normalized for the same apparent disk temperature of the Sun.

distribution, the latter was smoothed by the aerial pattern of the 32-element interferometer. The smoothed curves for scanning directions of 90° and 64° with respect to the central meridian of the Sun are in fair agreement with our one-dimensional distributions when the Sun was scanned in the same directions. The results are shown in Figure 7.

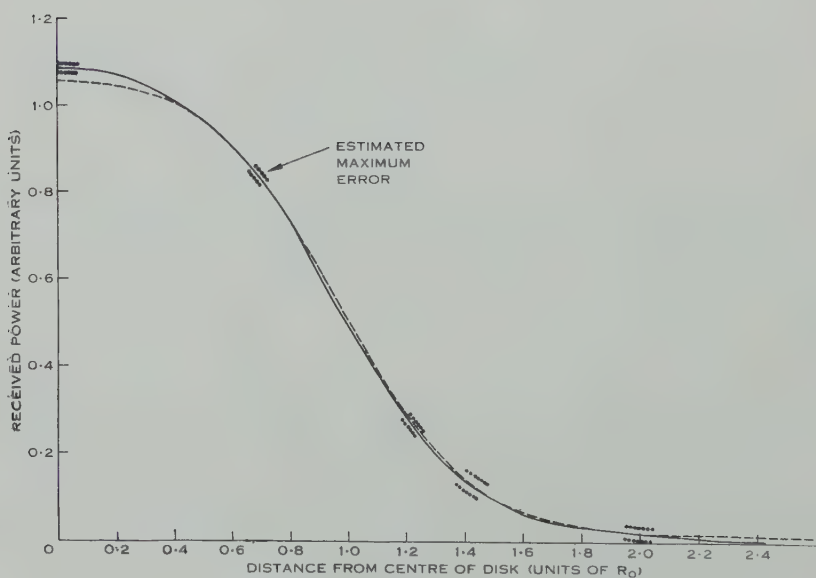
The remaining discrepancy we believe to be greater than our errors of observation. Since both series of observations are subject to limitations of resolving power associated with the aerial aperture or spacings employed, we examined the distribution of the Fourier components of the two distributions. The distribution for the equatorial scanning direction was used in each case. Our distribution shows significant components nearly to the cut-off frequency corresponding to the size of our aerial system (aperture 350 wavelengths). O'Brien and Tandberg-Hanssen used spacings up to 260 wavelengths, but were unable to detect any components greater than that corresponding to a spacing of 200 wavelengths. In addition, on July 6 and 7 when O'Brien and Tandberg-Hanssen obtained one of the records, there was a localized bright region of low intensity (see Fig. 1 (b)) which would not be detected in their experiment and would have produced a small error in the results.

(c) Comparison with Eclipse Results

We have shown above that Stanier's observations do not agree with the present ones. This may merely reflect actual changes in the quiet Sun, as seen at a wavelength of 60 cm, between the two observing periods. To examine



(a)



(b)

Fig. 7.—Comparison of our one-dimensional distributions (full line) with those derived on smoothing the two-dimensional distribution of O'Brien and Tandberg-Hanssen with our aerial beam (broken line). (a) Scanning direction 90° , (b) scanning direction 64° .

this possibility various brightness distributions were compared with eclipse results.

O'Brien (1953) found reasonable agreement of Stanier's distribution with the 1948 eclipse results of Christiansen, Yabsley, and Mills (1949). But agreement with an eclipse curve is not very significant, though a disagreement is, because the eclipse curves are insensitive to widely different brightness distributions; for instance, Blum (1953) showed that an elliptical model of the Sun could also explain the eclipse results of Christiansen, Yabsley, and Mills.

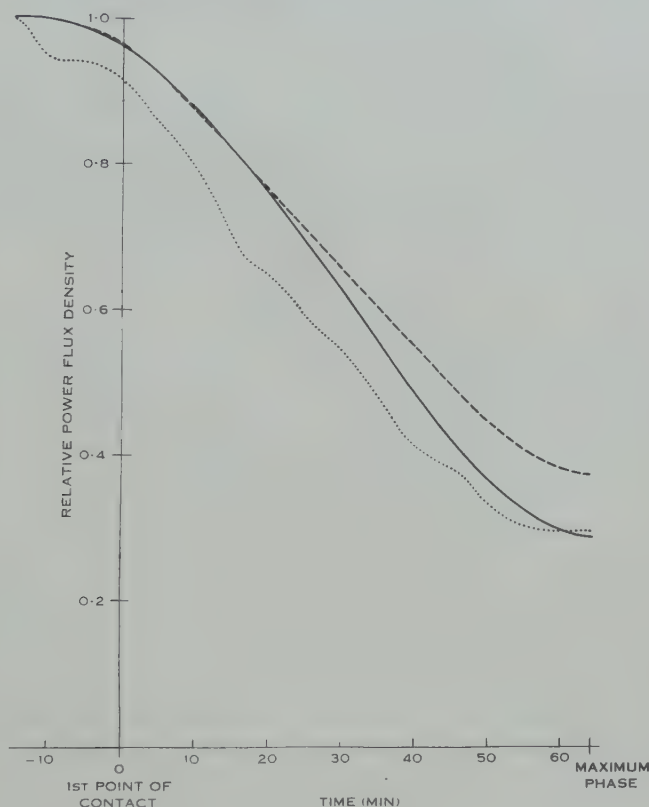


Fig. 8.—Comparison of 1948 eclipse curves at 50 cm with the artificial eclipse curve derived for the brightness distribution of O'Brien and Tandberg-Hanssen at 60 cm.

- Artificial eclipse curve derived for the distribution of O'Brien and Tandberg-Hanssen.
- Observed curve for 1948 eclipse.
- Curve derived for the quiet Sun by Christiansen, Yabsley, and Mills.

However, the 1948 eclipse curve derived for the quiet Sun by Christiansen, Yabsley, and Mills for the observing station, Strahan, does not agree with the curve obtained by artificially eclipsing the distribution of O'Brien and Tandberg-Hanssen (see Fig. 8). It is true that the 1948 quiet Sun curve was derived

indirectly and may not be correct. However, at the maximum phase of the eclipse the artificial eclipse curve gives the same residual as the directly observed experimental curve. The former refers to the quiet Sun while the latter clearly exhibits the presence of excess radiation from localized disturbed regions situated in the eclipsed part of the Sun. Only the north polar region of the Sun was not eclipsed. It is difficult to reconcile the presence of the same residual unless the quiet Sun distribution has changed during the period from 1948 to 1954. The alternative explanation is that at the time of the 1948 eclipse the radiation from localized regions in the vicinity of the pole (that is, the uneclipsed part) was equal to or greater than that from these regions in the equatorial zone of the Sun. This would seem most unlikely.

O'Brien and Tandberg-Hanssen (1955) claim that their results conform with 1954 eclipse observations. Their radio brightness distribution also agrees with our observations. Therefore the disagreement with the 1948 eclipse curve suggests that the quiet Sun may have changed during the period. Such a change might be expected, since it is well known that marked changes in the appearance of the solar corona occur during the solar cycle.

V. CONCLUSIONS

One-dimensional distributions over the quiet Sun have been obtained for the present minimum phase of the solar cycle. The method employed draws a clear distinction between the quiet Sun radiation and that from temporary and localized disturbed regions. This is not so for the method employed by Stanier who derived the brightness distribution over the quiet Sun at the same wavelength during the last maximum phase of the solar cycle. The measurements described here do not agree with Stanier's results. However, the measurements are in reasonably good agreement with the two-dimensional distribution over the Sun at the same wavelength derived by O'Brien and Tandberg-Hanssen with Stanier's method. Evidence is presented which suggests that the brightness distribution over the quiet Sun may have changed between the last sunspot maximum and the present sunspot minimum. But it is not possible to say whether this change fully accounts for the discrepancy between Stanier's and our results, or whether Stanier's quiet Sun distribution was affected by radiation from disturbed regions which might have been present in the Sun at the time of observations.

VI. ACKNOWLEDGMENTS

The authors are indebted to Mr. C. Fryar for assistance in the observational work; Dr. W. N. Christiansen, Mr. J. A. Warburton, and Mr. S. F. Smerd for valuable discussions; and Dr. J. L. Pawsey for helpful advice and criticism throughout the period of the project. During this investigation both the authors were holders of Fellowships awarded by the Australian Government under the Colombo Plan.

VII. REFERENCES

- ALON, I., ARSAC, J., and STEINBERG, J. L. (1953).—*C.R. Acad. Sci., Paris* **237**: 300.
ALON, I., ARSAC, J., and STEINBERG, J. L. (1955).—*C.R. Acad. Sci., Paris* **240**: 595.
BLUM, E. J. (1953).—*C.R. Acad. Sci., Paris* **237**: 135.

- BRACEWELL, R. N. (1955).—*Aust. J. Phys.* **8**: 200.
- CHRISTIANSEN, W. N., and WARBURTON, J. A. (1953*a*).—*Aust. J. Phys.* **6**: 190.
- CHRISTIANSEN, W. N., and WARBURTON, J. A. (1953*b*).—*Aust. J. Phys.* **6**: 262.
- CHRISTIANSEN, W. N., and WARBURTON, J. A. (1955).—*Aust. J. Phys.* **8**: 474.
- CHRISTIANSEN, W. N., YABSLEY, D. E., and MILLS, B. Y. (1949).—*Aust. J. Sci. Res. A* **2**: 506.
- COVINGTON, A. E., and BROTEN, N. W. (1954).—*Astrophys. J.* **119**: 569.
- O'BRIEN, P. A. (1953).—*Mon. Not. R. Astr. Soc.* **113**: 597.
- O'BRIEN, P. A., and TANDBERG-HANSEN, E. (1955).—*Observatory* **75** (884): 11.
- STANIER, H. M. (1950).—*Nature* **165**: 354.
- SWARUP, G., and PARTHASARATHY, R. (1955).—*Observatory* **75** (884): 8.

APPARENT INTENSITY VARIATIONS OF THE RADIO SOURCE HYDRA-A

By O. B. SLEE*

[*Manuscript received August 15, 1955*]

Summary

This paper describes the results of a detailed study of the flux density of the discrete radio source, Hydra-A. Over a period of 12 months, approximately 200 measurements of the Hydra-A intensity were compared with a similar number of observations on other strong sources. It was concluded that the observed flux density of Hydra-A is much more variable than that of these other sources, although no periodic changes have yet been detected. Possible mechanisms for the observed intensity changes are discussed.

I. INTRODUCTION

In the course of a survey of cosmic radio noise in 1946 Hey, Parsons, and Phillips noticed that the intensity recorded from a region in Cygnus exhibited marked variability; they suggested that this variable radiation originated in a source of small angular size. In 1947 Bolton and Stanley showed that such a discrete source of radiation did exist, and in addition were able to detect a few similar objects. Since these initial discoveries, several hundred discrete radio sources have been reported by different workers in both northern and southern hemispheres. It was soon demonstrated that the intensity fluctuations which led to the discovery of the Cygnus source were ionospheric in origin and that the emission of the source was constant to within the accuracy of measurement. However, the interest in the constancy of other discrete sources has remained because of the great astronomical significance of any variability which can be established.

From the results of a survey in which the intensities of about 100 sources with declinations between 10° and 80° were measured almost daily for 18 months, Ryle and Elsmore (1951) concluded that none showed any variation in intensity greater than 10 per cent. No such systematic survey has yet been undertaken by workers in the southern hemisphere, but from the many observations made at this laboratory on the stronger sources it was concluded that their intensities appeared to be constant within 10 per cent., a figure set by the errors of measurement. However, during the course of a discrete source survey (Bolton, Stanley, and Slee 1954), the author was impressed by apparent changes in the flux density of the source Hydra-A. The position given by Mills (1952) is R.A. 09 hr 15 min 46 sec, Dec. $-11^\circ 55'$ (epoch 1950) and its I.A.U. number 09S1A. The intensities for two consecutive records differed by 30 per cent., although

* Division of Radiophysics, C.S.I.R.O., University Grounds, Sydney.

sources which appeared on the records at earlier and later rising times showed no significant change. In April 1954, the 110 Mc/s sea interferometer at Dover Heights, Sydney, became available for a systematic study of this source, which was recorded as often as possible until December 1954 when it was decided to transfer the observations to a meridian transit instrument on a site near Sydney ; by this means it was hoped to eliminate effects associated with very low angles of observation.

Since the measurement of absolute flux density is subject to factors which may vary in an unknown manner with time, the observed intensity of the Hydra source on any day was compared with that of one or more of the bright radio sources in Taurus, Virgo, Centaurus, and Fornax, all of which were observed within a few hours of each other. In this way, measurements of the intensity of the Hydra source were largely freed from the effects of calibration errors, and, in addition, a series of records was obtained of other sources for comparison with the suspected variable source.

The observations support the conclusion that the observed intensity of Hydra-A is more variable than that of the four comparison sources.

II. THE OBSERVATIONS

(a) *Rising Measurements*

The equipment and techniques used in this series of observations have been described in detail elsewhere (Bolton and Slee 1953). The sensitivity of the equipment was monitored by frequent observations of one or more of the intense sources in Taurus, Virgo, Centaurus, and Fornax. Information concerning the five sources is given in Table 1. Two examples of records obtained

TABLE 1
DETAILS OF THE FIVE SOURCES OBSERVED

Source		Flux Density* at 100 Mc/s ($\text{W m}^{-2} (\text{c/s})^{-1} \times 10^{24}$)	Approximate Position (1950)		Rising Coordinates		No. of Observations	
Constel- lation	I.A.U. No.		R.A. (hr min)	Dec.	Sidereal Time (hr min)	Azimuth	Rising	Transit
Hydra ..	09S1A	2.5	09 16	$-11^{\circ} 55'$	02 45	105°	108	91
Taurus ..	05N2A	19.0	05 31	$22^{\circ} 00'$	00 30	60°	33	65
Virgo ..	12N1A	12.0	12 28	$12^{\circ} 40'$	07 00	75°	25	95
Centaurus	13S4A	17.0	13 22	$-42^{\circ} 40'$	04 45	140°	20	57
Fornax ..	03S3A	2.5	03 20	$-37^{\circ} 18'$	19 10	135°	19	—

* The intensities quoted refer to interferometer measurements at 100 Mc/s.

in this series are shown in Figure 1, in which Hydra-A, rising at 02 hr 45 min, is in both cases calibrated by the Taurus source rising at 00 hr 30 min. These records illustrate a typical change of 20 per cent. in the Hydra-A flux density between the two observations, while no significant change is visible in the

Taurus source; it is also apparent that the difference in the intensities of Hydra-A was maintained for the period of observation lasting about 1 hr.

It is believed that scintillation effects, examples of which appear on the interference patterns of Figure 1, are the major source of error in reading the records. In order to ascertain if scintillations caused systematic errors in the recorded intensities of the sources, the scintillation indices for the records of

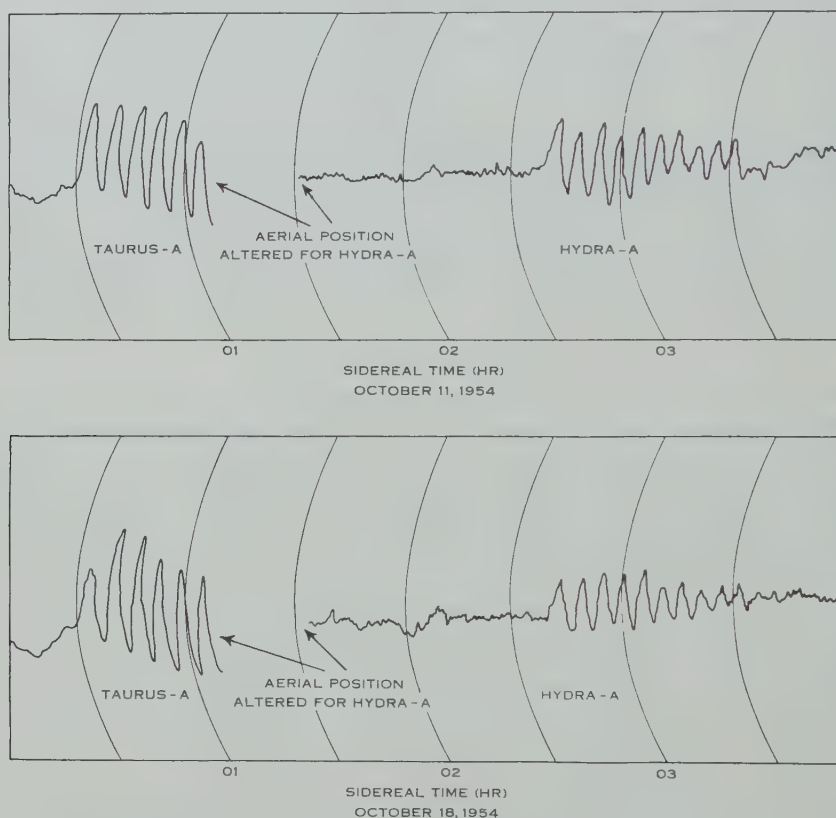


Fig. 1.—Typical records of Hydra-A and a comparison source obtained in the rising observations. The average deflection due to Hydra-A is 20 per cent. greater in the upper record than in the lower. Scintillations are visible on the interference fringes of both records.

the Hydra source were plotted against the estimated daily deviations from the mean intensity. The resulting plot, shown in Figure 2, indicates that there is no direct correlation between the apparent intensity of the source and the scintillation index.

(b) Transit Observations

From early December 1954 to July 1955, the Hydra source, together with the three comparison sources in Taurus, Virgo, and Centaurus, was recorded

almost daily during meridian transit, using portion of the 85 Mc/s "cross" aerial, of which a description by Mills, Little, and Sheridan is in preparation.

The east-west arm of the aerial produces a fan-shaped response pattern, 0.6° east-west and 50° north-south, between half-power points. A large number of radio sources at different declinations therefore passes through the beam at meridian transit. By amplifying the output of the east-west arm and its associated receiver in a D.C. amplifier, Hydra-A and the three comparison sources were recorded at transit. Two sample records taken on successive nights

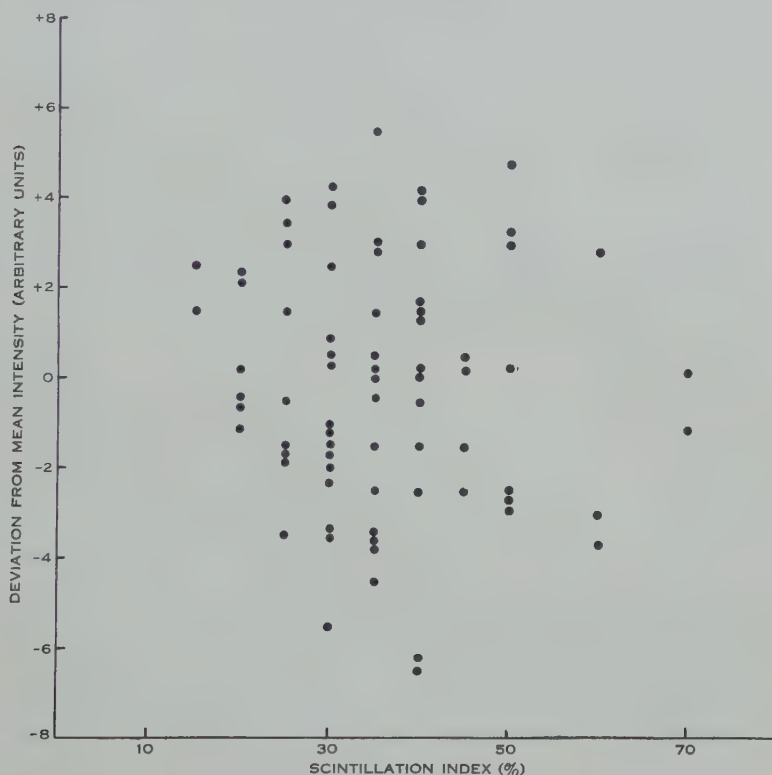


Fig. 2.—A correlation diagram of the scintillation indices and estimated deviations from the mean intensity for the rising series of observations on Hydra-A.

are shown in Figure 3, in which the deflection due to Hydra-A at 09 hr 16 min has changed by 30 per cent. ; other sources on the records show no appreciable alteration.

Any sensitivity changes which took place during the series were assumed to be very slow in character, and were measured by computing running intensity means of 10 consecutive records of each of the comparison sources ; in practice it was found that only small corrections were needed to bring the intensity measurements to a common scale for direct comparison with each other over long periods of time.

III. ANALYSIS OF RESULTS

The daily measured intensities for Hydra-A and a comparison source are plotted in Figure 4. The comparison source plotted for the rising measurements is a composite one, obtained by bringing the intensity measurements of the Taurus, Virgo, Fornax, and Centaurus sources to a common scale; this procedure was necessary as normally only one comparison source was recorded at any particular time of the year. In the transit observations, Virgo-A has been shown for comparison, but similar plots are obtained for the Taurus and Centaurus sources.

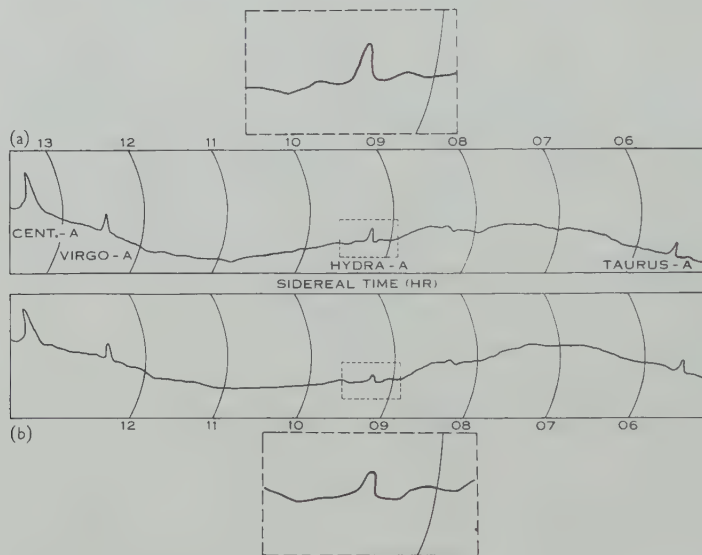


Fig. 3.—Typical records of Hydra-A and the three comparison sources in Taurus, Virgo, and Centaurus obtained during the transit observations. In the upper record of April 4, 1955, the intensity of Hydra-A is 10 per cent. above the mean and in the lower record of April 5, 1955, 30 per cent. below the mean.

It is obvious from Figure 4 that the observed intensity of Hydra-A possesses a greater range of values than the comparison sources. This difference is probably shown more effectively in Figures 5 (a) and 5 (b), which are histograms of the intensity distributions for the rising and transit observations respectively.

In interpreting the significance of the differences between Hydra-A and the comparison sources, it should be noticed that in both series of observations the suspected variable was compared with sources which were of higher signal-to-noise ratio at the receiver, except for the recordings of Fornax-A in the rising observations and Taurus-A in the transit series. One might therefore expect a larger scatter in the Hydra-A intensities. However, when the transit results were subjected to statistical analysis, details of which are given in Table 2, it was found that the standard deviations of the three comparison sources were very similar despite large differences in signal-to-receiver noise fluctuation ratios.

It may be concluded that the noise fluctuation level was not high enough to seriously affect the accuracy of intensity measurements for any of the sources.

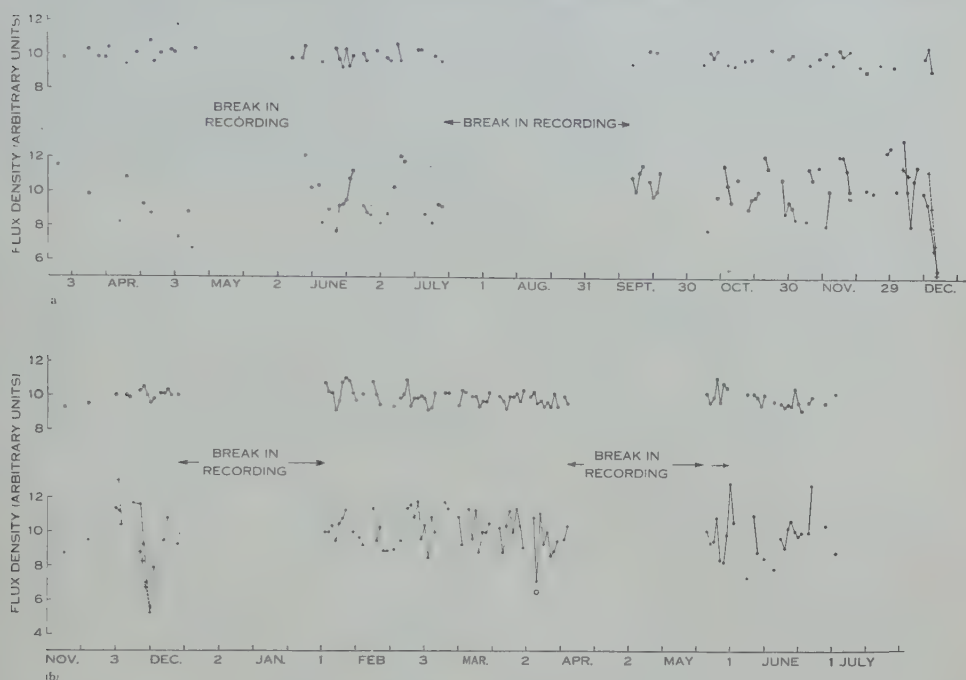


Fig. 4.—Plots of the daily observed intensities of Hydra-A and the comparison sources. All sources have been brought to a common arbitrary intensity scale of mean value 10 units. (a) and (b) refer to the rising and transit observations respectively. The upper diagram in each figure shows the results for the comparison source, and the lower for Hydra-A. The open circle of (b) refers to an interferometer measurement. The crosses plotted in December represent values transferred from the other series of measurements during the overlap of observations.

The above evidence is suggestive of a real variation in the apparent intensity of Hydra-A, but it is insufficient for a conclusive demonstration. Fortunately,

TABLE 2
RELATIONSHIP BETWEEN SIGNAL-TO-NOISE RATIO AND
ACCURACY OF INTENSITY MEASUREMENTS

Source	Signal-to-noise Ratio	% Standard Deviation
Hydra-A	10	12.2
Virgo-A	15	5.3
Taurus-A	10	6.0
Centaurus-A ..	30	6.1

simultaneous or nearly simultaneous observations of the source with two different instruments are available. These serve to strengthen considerably the belief in a real variation.

Large intensity fluctuations of the Hydra source have, to some extent, been confirmed by Carter (1955), who observed the source at transit with an east-west interferometer operating at 101 Mc/s with a spacing of either 90 or 1000 wavelengths. The observations were made during April, May, June 1955, and on some of these days simultaneous records were obtained by the author with the east-west arm of the "cross". A statistical comparison of the Hydra-A intensities measured on the two systems showed a significant correlation. Three very significant reductions in the Hydra-A flux density,

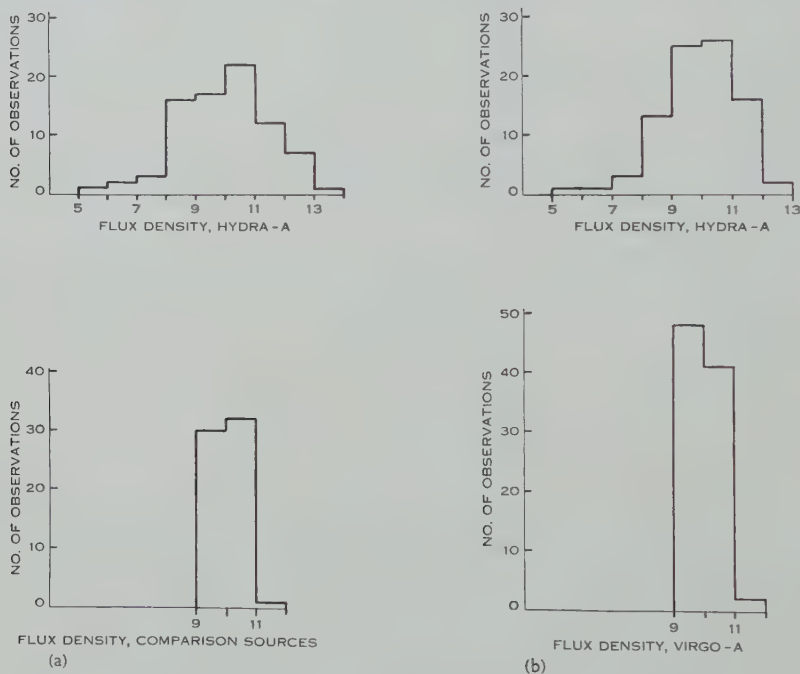


Fig. 5.—Histograms showing the distributions of intensities for Hydra-A and the comparison sources. (a) and (b) refer to the rising and transit observations respectively.

amounting to as much as 70 per cent. of the mean intensity, were measured with the interferometer during this period, and on one of these occasions, April 5, 1955, a simultaneous record was obtained by the author; the two values of flux density, shown plotted in Figure 4, were well correlated with each other on this particular day. It should be pointed out that, on the three days concerned, the intensity of the source appeared constant during the interferometer observations lasting about 1 hr.

During the transition from rising to transit observations in December 1954, records were obtained of Hydra-A with both instruments on six days, the times of observation differing by approximately 6 hr. Both sets of intensity measurements for the days concerned have been superimposed in Figure 4, which shows a high correlation between the two series.

The flux density of Hydra-A does not appear to change in a periodic manner over an interval of a few days, as is the case for many variable stars, but rather varies randomly about the mean from day to day. This conclusion is supported by the results of an autocorrelation test of the transit measurements. Interferometer observations have shown that the source intensity is sensibly constant for periods approaching 6 hr; the average "fading time" of the source therefore appears to lie between 6 and 24 hr.

IV. DISCUSSION

The observations support the conclusion that the observed flux density of the discrete radio source in Hydra is variable. Possible mechanisms for the production of the observed intensity variations may operate either in the source or during transmission of the radiation through space or the terrestrial atmosphere. No definite information is as yet available to fix the origin of the Hydra-A radiation, but obviously if this can be done, then possible mechanisms may be more clearly defined.

Using the full resolution of the 85 Mc/s "cross", Mills (unpublished data) finds that the position of the Hydra source is R.A. 09 hr 15 min 40 sec ± 4 sec, Dec. $-11^{\circ} 52' 5'' \pm 2''$ (1950). The radio position is practically coincident with that of a faint galaxy which has been photographed by Baade, and its spectrum obtained by Minkowski using the 200 in. telescope (personal communication, Minkowski to Mills). It is found to display a double nucleus suggesting a very close approach of two galaxies, but there are no abnormalities in the spectrum to suggest that an actual collision may be in progress as in the case of the Cygnus radio source. It is not possible therefore to positively identify the radio source with the galaxy. However, if the optical object is responsible for the Hydra-A radiation, it is unlikely that the observed short-term variations in the radio flux density could be due to inherent changes in the radio emission of the galaxy unless the radio energy is emitted from only a small region. Measurements by Carter (1955) suggest that the half-brightness size of the Hydra source in the east-west direction is approximately $1.5'$, a figure which is consistent with the dimensions of the galaxy.

Short-period fluctuations or scintillations in radio-source intensities are known to occur quite frequently, and have been studied extensively by several workers. Scintillation activity is believed due to the diffraction or focusing of the radiation by irregularities of about 5 km dimensions in the F and E regions of the ionosphere. However, very few cases of general absorption of the radiation at frequencies of the order of 100 Mc/s are known. Bolton, Slee, and Stanley (1953) reported that on three occasions out of about 1000 observations of the Cygnus source at rising, marked general absorption for periods up to 1 hr were noticed, but on these occasions fast and intense scintillations were also present. Since the flux density variations of Hydra-A show no correlation with scintillation activity, it is very unlikely that they are produced by the same mechanisms. The characteristics of the E and F regions of the ionosphere were examined at approximately the times of the observed intensity extremes shown

by the Hydra source; the hourly values recorded at the Christchurch (New Zealand) ionospheric station were compared with the rising measurements, and the Canberra (Australia) hourly values used for the transit observations. No significant departures from normality were recorded in the critical frequencies of the ionospheric layers on these occasions. The daily magnetic variations recorded at Toolangi (Australia) for the same days were also examined with negative results. It must therefore be concluded that the large reductions in the Hydra-A intensity, which are believed to have special significance, cannot be explained by any well-known ionospheric effect. However, in view of the small angular size of the source compared with the comparison sources, it is possible that ionospheric effects may be more severe. One method of estimating the effect of the ionosphere on the radiation from sources of different angular sizes is to study their average scintillation indices. The radiation from the Hydra-A and Virgo-A sources at rising traverses equal path lengths through the ionosphere, and the sources differ in angular size by a factor of three. Since the average scintillation indices for the two sources over a period of 12 months are almost equal, it seems improbable that an ionospheric mechanism, selective to source size, produces the observed intensity changes in Hydra-A.

It has been suggested that a potential mechanism for the production of intensity variations in the Hydra source is the Faraday effect, that is, the rotation of the plane of polarization of a plane polarized wave during its passage through an ionized gas in the presence of a magnetic field. Murray and Hargreaves (1954) showed that such a mechanism was probably responsible for the slow fading of lunar echoes at a frequency of 120 Mc/s during transmission through the terrestrial ionosphere. However, in the case of the Hydra-A radiation, the ionized region responsible need not necessarily be located near the Earth, but could conceivably exist near the source or in interstellar space. In order to account for the observed intensity variations, the Hydra-A radiation would need to be plane polarized to the extent of at least 50 per cent. If present, polarization percentages of this order should be relatively easy to detect.

Finally, it is interesting to consider the implications of assuming that the variability of intensity is an intrinsic property of the source. It is clear that the physical dimensions of the source would be unlikely to exceed the distance an electromagnetic wave is propagated in a time of the order of the "fading time", that is, a distance of one "light-day". Taken in conjunction with the angular size of about $1.5'$, this would suggest that the source must be within a few parsecs of the Sun. Such a distance, if correct, would seriously undermine some of the current ideas about the distribution of radio sources.

V. ACKNOWLEDGMENTS

The author is indebted to Mr. R. X. McGee of this laboratory for much of the observational work connected with the investigation and for helpful discussion of the results. Thanks are also due to Mr. B. Y. Mills and Mr. A. W. Carter for information in advance of publication on the position and angular size respectively of the Hydra radio source.

VI. REFERENCES

- BOLTON, J. G., and SLEE, O. B. (1953).—*Aust. J. Phys.* **6** : 420.
BOLTON, J. G., SLEE, O. B., and STANLEY, G. J. (1953).—*Aust. J. Phys.* **6** : 434.
BOLTON, J. G., STANLEY, G. J., and SLEE, O. B. (1954).—*Aust. J. Phys.* **7** : 110.
CARTER, A. W. L. (1955).—*Aust. J. Phys.* **8** : 564.
MILLS, B. Y. (1952).—*Aust. J. Sci. Res. A* **5** : 266.
MURRAY, W. A. S., and HARGREAVES, J. K. (1954).—*Nature* **173** : 944.
RYLE, M., and ELSMORE, B. (1951).—*Nature* **168** : 555.

ROTATION AND OTHER MOTIONS OF THE MAGELLANIC CLOUDS FROM RADIO OBSERVATIONS

By F. J. KERR* and G. DE VAUCOULEURS†

[*Manuscript received August 15, 1955*]

Summary

The radial velocities measured in a survey of the 21 cm line from the Magellanic Clouds are discussed, principally in terms of the median velocity for each line profile. It is shown that each Cloud is rotating, and the analysis confirms the view, based on optical evidence, that the Clouds are flattened and tilted systems. The gas is probably the least flattened constituent of each Cloud and shows a large decrease of rotational speed with increasing distance from the equatorial plane.

The mean radial velocities of the Large and Small Clouds with respect to the Sun, weighted according to the 21 cm surface brightness, are found to be +280 and +161 km/sec. Other internal motions, both systematic and random, are also indicated by the profiles.

I. INTRODUCTION

New possibilities for studying large-scale motions of galactic systems have been opened up by the detection of 21 cm line radiation from interstellar hydrogen. This radiation, in common with other types of radio-frequency radiation, is able to penetrate through very large distances in interstellar space. In addition, it has made possible for the first time the measurement of velocities in these extensive regions which are visible to the radio waves but not to light.

Studies of the line radiation suffer from the limited resolution which is a characteristic of radio equipment. This is not always a disadvantage, however, when systematic effects are being investigated. For example, when velocity measurements are made with limited resolution, an average is directly obtained over a substantial area. On the other hand, when optical observations are made on individual objects (stars, emission nebulae, etc.), a large sample is required to eliminate the effects of the velocity dispersion among such objects.

The Magellanic Clouds,‡ being the nearest external galaxies, are specially suitable objects for study, since each subtends an angle of several degrees and can therefore be explored in some detail. A preliminary survey of 21 cm radiation from the Clouds has been described in an earlier paper (Kerr, Hindman, and Robinson (1954), referred to below as paper I), which contains an account of the technique employed and the observational results obtained, together with a partial analysis. Despite the nearness of the Clouds, less is known about

* Division of Radiophysics, C.S.I.R.O., University Grounds, Sydney.

† Yale-Columbia Southern Station, Mt. Stromlo, Canberra.

‡ A review of current knowledge of the Clouds has recently been published by Buscombe, Gascoigne, and de Vaucouleurs (1954).

their motions than is the case for several other external systems. This is largely due to their southern location, which makes them inaccessible to the few optical telescopes equipped with suitable spectrographs. The optical information so far available is restricted to radial velocities for 17 emission nebulae in the Large Cloud and one in the Small Cloud, obtained some 40 years ago during the D. O. Mills Expedition of the Lick Observatory to the southern hemisphere (Wilson 1917).

Following a suggestion made by Wilson, Hertzsprung (1920) re-examined the data for evidence of rotation of the Large Cloud. He found no clear rotation, but concluded that the observations could be interpreted as indicating a common translational motion of the two Clouds through space, nearly at right angles to the line of sight. This suggestion was also discussed by Luyten (1928), and later extended by Wilson (1944), who derived a space velocity of 550 km/sec towards $l=140^\circ$, $b=-32^\circ$. Such a large velocity is unlikely, however, in view of the small velocity dispersion among galaxies belonging to the local group (Mayall 1946).

In this paper, the distribution of the 21 cm velocities reported in paper I will be used to show that each Cloud is rotating and then the other motions observed in the Clouds will be briefly considered.* The problem of deriving the masses of the Clouds from observations of the rotational and random motions will be considered in a later paper.

The treatment followed here is largely a two-dimensional one, as the analysis was initially based on the median velocity of each profile, which is the simplest single parameter. It is clear, however, that the rotational motion is far more complicated than indicated by the simple treatment, and radio studies open up possibilities of deriving the three-dimensional distribution of gas density and rotational motion. Some information on the space distribution can be obtained from the present data, but a full three-dimensional study, involving detailed consideration of the fine structure of the line profiles, has been left until better observations are available.

II. OBSERVATIONAL DATA

The observations described in paper I were carried out using an aerial of beam width 1.5° between half-power points,† and a receiver with a bandwidth of 40 kc/s, which corresponds to a velocity range of 8 km/sec.

Profiles were obtained for about 250 different regions in the Clouds. Each profile gives the received power as a function of frequency, or velocity, and thus gives the velocity distribution of the radiating gas, averaged over a volume in the Cloud defined by the aerial beam.

The scanning technique employed in these observations was such that the profiles were built up from a number of constant-frequency sweeps across the

* A short discussion of the velocity data was included in a paper presented to a meeting of the American Astronomical Society (Kerr and Hindman 1953). The tentative interpretation given there, including a large translational motion, was necessarily based on a very brief perusal of the data, and this has since been considerably modified, in view of the new optical results giving evidence of spiral structure (de Vaucouleurs 1954a).

† The beam width was incorrectly quoted in paper I as 1.0° .

region containing the Clouds. Profiles were drawn for points on a lattice spaced 1° in declination and 10 min in Right Ascension. With a beam width of 1.5° , these points are substantially independent. Adjacent profiles are not quite independent, however, since these preliminary observations did not delineate each profile completely and in some cases points had to be estimated by interpolation from the shapes of neighbouring profiles.

Some typical profiles were shown in paper I, and others are presented below. In the central part of the Small Cloud the curves are smooth, but elsewhere they are more complex. The line widths (between half-power points) are about 50 km/sec, with minor variations over the region.

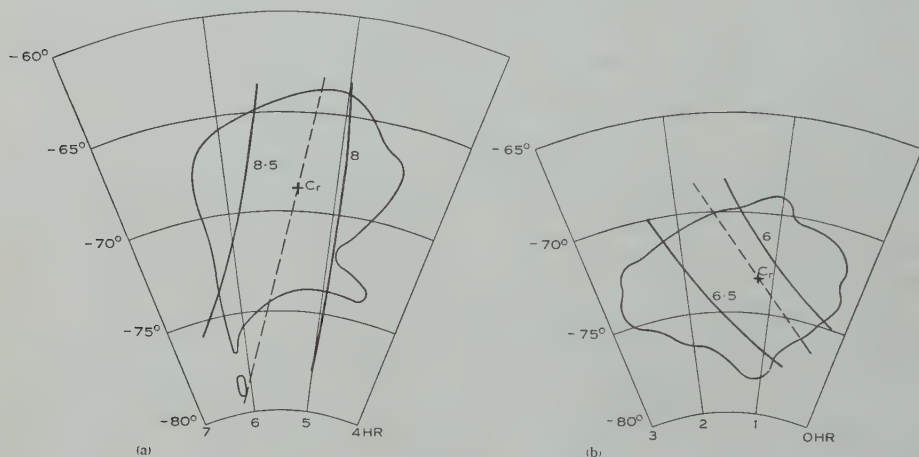


Fig. 1.—Galactic rotation correction. The contours (thick lines) show, in km/sec, the amount to be subtracted from the Magellanic Cloud residual velocities for a 10 km/sec increase in the adopted value of v_G , the galactic rotation velocity. The dashed lines show for comparison the major axes derived from the radio data. (a) The Large Cloud, (b) the Small Cloud.

In the main discussion in this paper, each profile is represented by a single parameter. This is the *median* radial velocity, the velocity where the ordinate halves the area under the profile.* In Section IV, however, the velocities of the major peaks in some of the Large Cloud profiles are discussed, and the significance of the median and peak velocities considered.

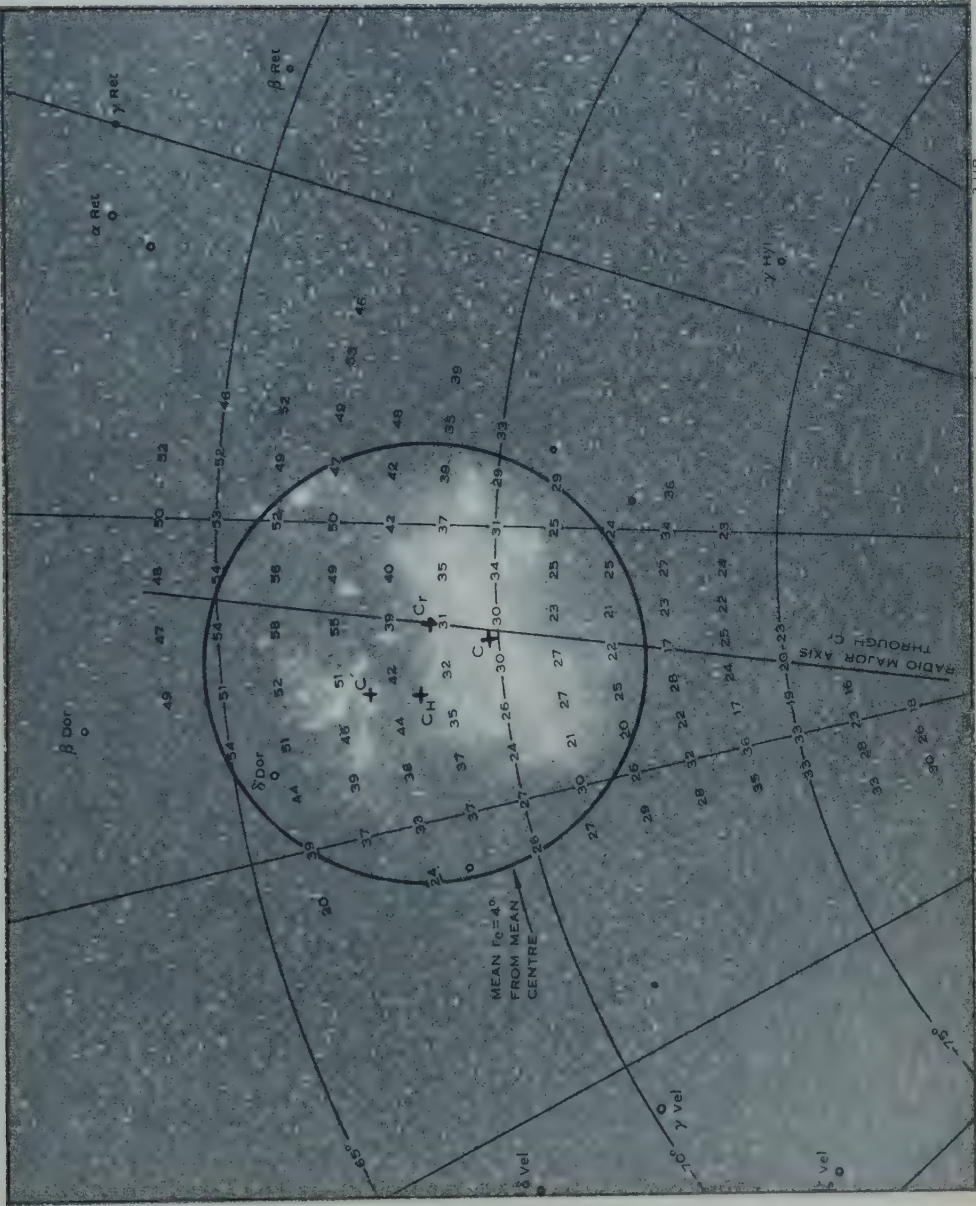
Smoothed contours of the observed median radial velocity were presented in Plate 2, Figure 3, of paper I. Before the results can be analysed, it is necessary to remove the effects arising from solar motion and galactic rotation. The standard values adopted in the reduction were as follows:

(a) Solar motion, 20 km/sec in a direction $\alpha = 18$ hr 00 min, $\delta = 30.0^\circ$ (1900) ($l = 23.5^\circ$, $b = 21.6^\circ$);

(b) galactic rotation, 270 km/sec towards $l = 57^\circ$, $b = 0^\circ$.

* This quantity was called the "mean radial velocity" in paper I, but it is more correctly a median.

MOTIONS OF THE MAGELLANIC CLOUDS



The Large Magellanic Cloud and the surrounding star field (Aero-Ektar photograph). On the transparency are shown the residual median radial velocities from the 21 cm profiles and a scale of coordinates (1950), together with the geometrical centre C , bright star centroid C' , H I centroid C'' , and centre of rotation C . The circle approximately encloses

region containing the Clouds. Profiles were drawn for points on a lattice spaced 1° in declination and 10 min in Right Ascension. With a beam width of 1.5° , these points are substantially independent. Adjacent profiles are not quite independent, however, since these preliminary observations did not delineate each profile completely and in some cases points had to be estimated by interpolation from the shapes of neighbouring profiles.

Some typical profiles were shown in paper I, and others are presented below. In the central part of the Small Cloud the curves are smooth, but elsewhere they are more complex. The line widths (between half-power points) are about 50 km/sec, with minor variations over the region.

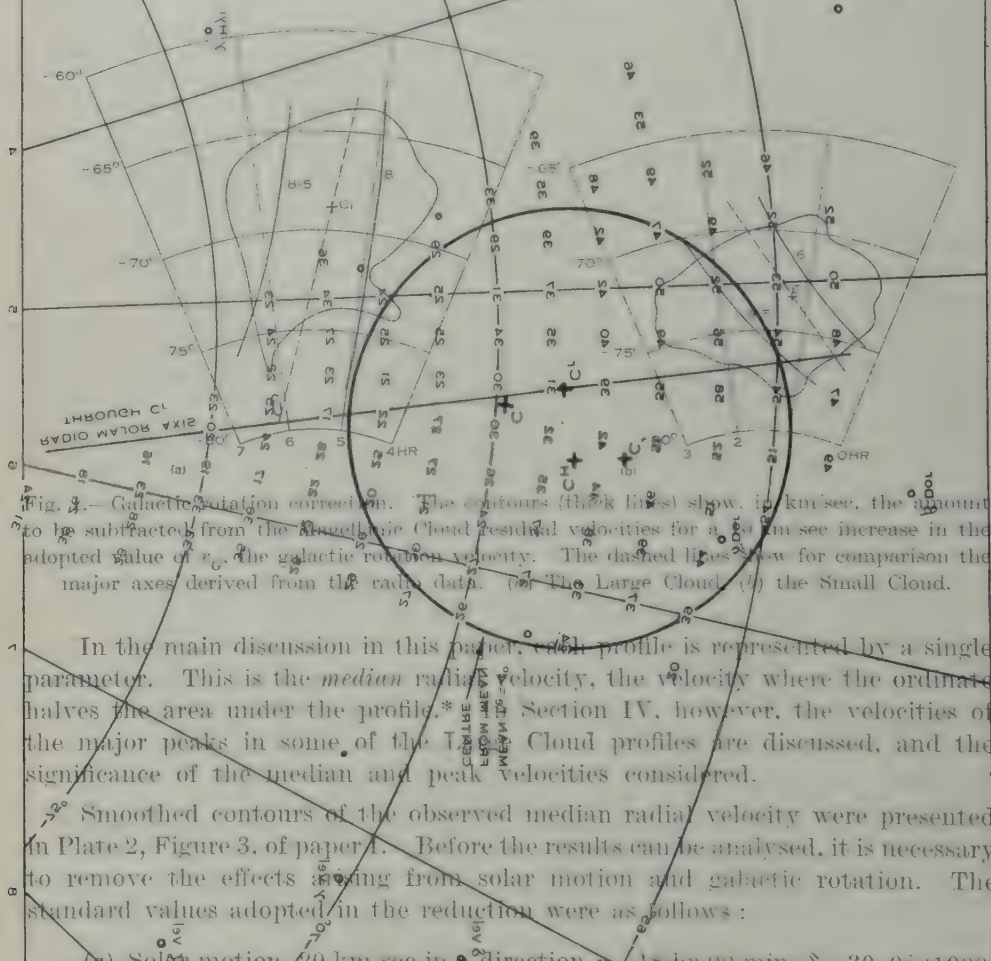


Fig. 3.—Galactic rotation correction. The contours (thick lines) show, in km/sec, the amount to be subtracted from the observed Cloud residual velocities for a 20 km/sec increase in the adopted value of v_0 , the galactic rotation velocity. The dashed lines show for comparison the major axes derived from the radio data. (a) The Large Cloud, (b) the Small Cloud.

In the main discussion in this paper, each profile is represented by a single parameter. This is the *median* radial velocity, the velocity where the ordinate halves the area under the profile.* Section IV, however, the velocities of the major peaks in some of the Large Cloud profiles are discussed, and the significance of the median and peak velocities considered.

Smoothed contours of the observed median radial velocity were presented in Plate 2, Figure 3, of paper I. Before the results can be analysed, it is necessary to remove the effects arising from solar motion and galactic rotation. The standard values adopted in the reduction were as follows:

- (a) Solar motion, 20 km/sec in a direction $\alpha = 18^h 00^m$, $\delta = 20^\circ 0'$ (1900) ($l = 23.5^\circ$, $b = 21.6^\circ$);
- (b) galactic rotation, 270 km/sec towards $l = 57^\circ$, $b = 0^\circ$.

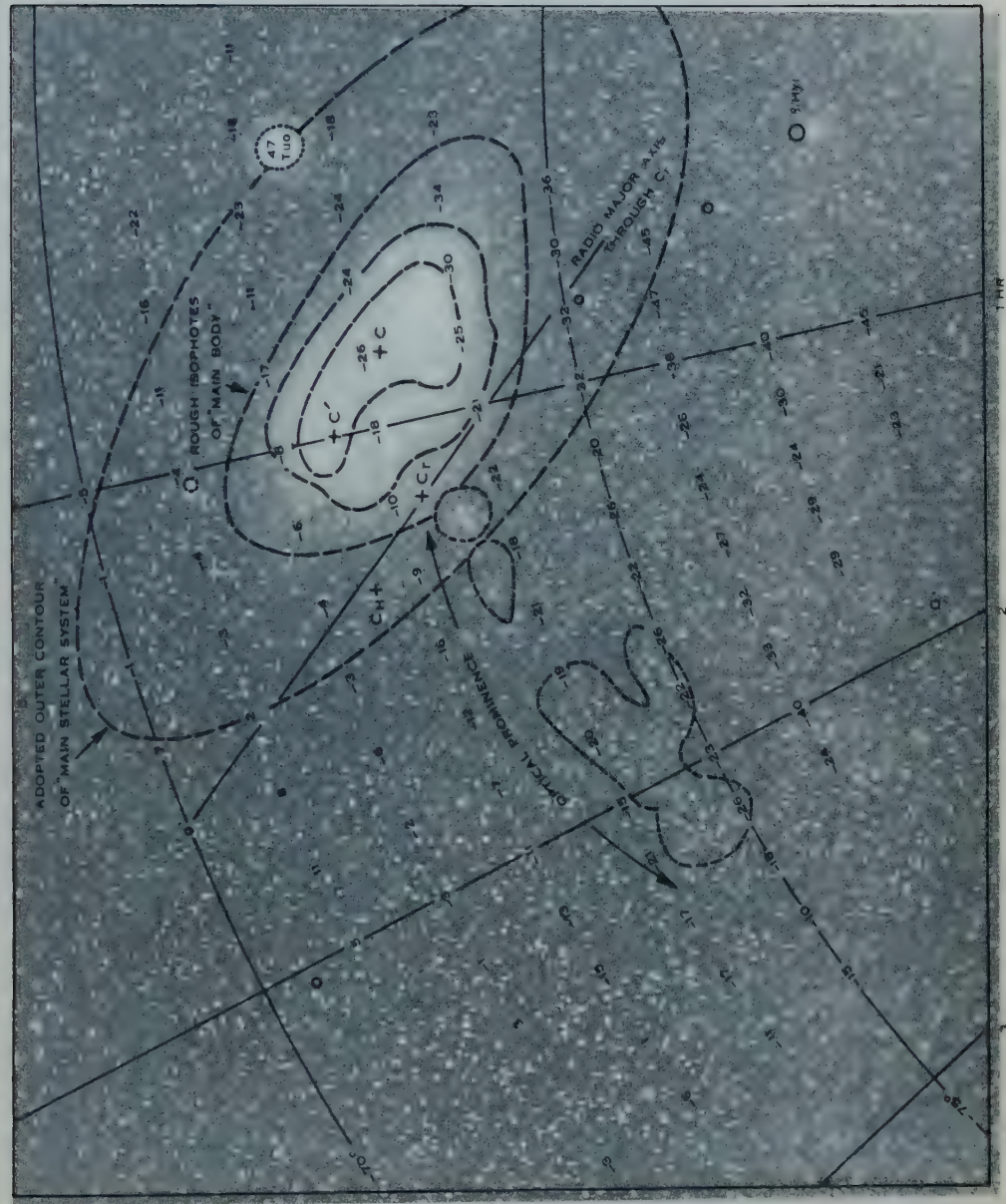
* This quantity was called the "mean radial velocity" in paper I, but it is more correctly a median.

MOTIONS OF THE MAGELLANIC CLOUDS



The Large Magellanic Cloud and the surrounding star field (Aero-Ektar photograph). On the transparency are shown the residual median radial velocities from the 21 cm profiles and a scale of coordinates (1950), together with the geometrical centre C , bright star centroid C' , $H\text{ I}$ centroid C'' , and centre of rotation C_0 . The circle approximately encloses

MOTIONS OF THE MAGELLANIC CLOUDS



The Small Magellanic Cloud and the surrounding star field (Aero-Ektar photograph). On the transparency are shown the residual median radial velocities, a set of coordinates, and the various centroids.

MOTIONS OF THE MAGELLANIC CLOUDS



The Small Magellanic Cloud and the surrounding star field (Aero-Ektar photograph). On the transparency are shown the residual median radial velocities, a scale of coordinates, and the various centroids.

MOTIONS OF THE MAGELLANIC CLOUDS



NGC 55: an asymmetrical spiral of the Magellanic type seen edgewise. (Reynolds reflector photograph, Commonwealth Observatory.)

The solar motion is so small that the exact choice of the velocity vector to be used is of little consequence. The value to be taken for galactic rotation, however, requires more consideration. Adoption of any other value than that used would change the absolute magnitude of all the residual velocities, but would have a smaller effect on conclusions based on the variation of velocity across each Cloud. The magnitude of such effects can be seen from Figure 1, in which is plotted the amount by which the residual velocity must be reduced for each increase of 10 km/sec in the galactic rotation velocity.*

The lines of equal correction are approximately parallel to the major axes used later in deriving rotational curves for the two Clouds. Hence the possibility of error in the adopted velocity of galactic rotation need not be considered when discussing the rotation of the Clouds, but is important in the derivation of the velocity of one Cloud relative to the other.

The residual median radial velocities, obtained by correcting the observed velocities for solar motion and galactic rotation, are given in Plates 1 and 2, superposed on photographs of the Clouds. Positive radial velocities are directed away from the observer. The accuracy of these individual velocities is discussed in Section III (c).

III. ROTATION CURVES FROM MEDIAN RADIO VELOCITIES

The Magellanic Clouds have the appearance of irregular systems. Their extent in depth, i.e. whether they are flattened or spheroidal, has long been uncertain. One of us (de Vaucouleurs 1954*a*, 1954*b*, 1955*a*, 1955*b*) has recently shown that both Clouds have considerable spiral structure in their outlying parts, and that their outer optical isophotes are elliptical in shape. These results suggest that the Clouds are flattened systems inclined to the line of sight; this view is supported by the great flattening observed in NGC 55, which we believe to be a galaxy of Magellanic type, seen edge-on (Plate 3).†

TABLE 1
GEOMETRICAL DATA

	Angle of Tilt	Position Angle of Major Axis	Coordinates of Centre <i>C</i> (1950)
LMC ..	$65 \pm 5^\circ$	$160 \pm 5^\circ \ddagger$	05 hr 24 min, -69.8°
SMC ..	$30 \pm 2^\circ$	$45 \pm 2^\circ$	00 hr 51 min, -73.1°

‡ An earlier value, 165° , was used in the analysis shown in Figure 2 (a).

Table 1 gives the values derived from the optical study, together with estimates of the probable errors, for (a) the angle of tilt, i.e. the angle between the line of sight and the equatorial plane, (b) the orientation of the major axis of the elliptical projection, measured anticlockwise from the northern end of the hour circle through the centre of the Cloud, and (c) the "geometrical" centre *C* of each Cloud, taken as the centre of the axial bar.

* $10 \cos b \cos (l - l_0)$ km/sec, where l_0 is the longitude of the galactic centre.

† Additional evidence that NGC 55 is probably of Magellanic type is obtained from recent 3.5 m measurements (Mills 1955).

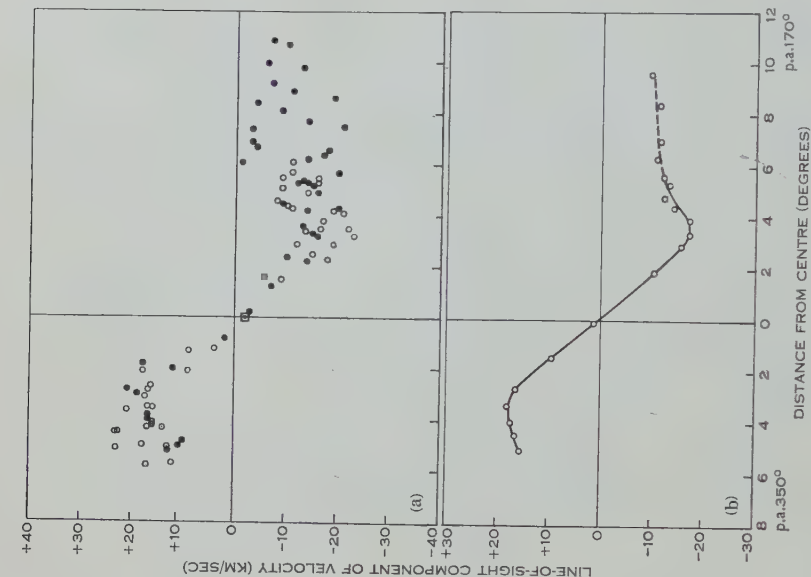


Fig. 2

Fig. 2.—Rotational curves for Large Cloud with respect to *optical centre*, C , 05 hr 24 min, -69.8° , position angle of major axis $= 165^\circ$ (axis of rotation $= 75^\circ$) using *median* radial velocities. Systemic velocity, $v_s = 30$ km/sec (corrected for solar motion and galactic rotation).

Fig. 3.—Rotational curves for the Large Cloud with respect to the derived *radio centre* of rotation, C_r , 05 hr 20 min, -68.8° , position angle of major axis $= 170^\circ$ (axis of rotation $= 80^\circ$) using *median* radial velocities. Systemic velocity, $v_s = 37$ km/sec.

(a) Individual points; ● major axis $\pm 20^\circ$, ○ sectors 20° – 60° from major axis, □ mean of points along minor axis.
(b) Mean curve; overlapping 9° point means.

(a) Large Cloud

Rotational motion can be looked for in either Cloud by plotting the residual velocities along its major axis. The analysis has been carried out in two stages.

In order to establish the existence of rotation in the Large Cloud, the optical values for the major axis and the centre were used and the median velocities plotted for points in a narrow sector within $\pm 20^\circ$ of the major axis. These points are shown by the filled circles in Figure 2 (a). They have a substantial scatter, but delineate a curve which is clearly of rotational type, analogous to those observed in M31 (Babcock 1939; Mayall 1950; Baade and Mayall 1951) and M33 (Mayall and Aller 1942).

After establishing the existence of rotational motion, the restriction to a narrow sector near the major axis was removed, and the whole set of data used to derive the best velocity curve. For the Large Cloud, all points within a sector $\pm 60^\circ$ about the major axis were treated in this way. In obtaining a rotational velocity from each observed radial velocity, allowance was made for the position angle of the point concerned and the tilt of the system, assuming circular orbits. The appropriate relations are given in Appendix I.

The first analysis of this type was done with reference to the optical centre *C*: the individual points are plotted in Figure 2 (a), and the mean curve in Figure 2 (b). There are maxima of the differential velocities at $3\text{--}4^\circ$ from the centre, on both sides of the axis, beyond which the velocities decrease in a way consistent with the Keplerian branch of a rotation curve.

The mean curve based on the optical centre of the Cloud is markedly asymmetrical. Since the Cloud itself is asymmetrical in form, we might conclude from the shape of this curve that the rotation pattern is also asymmetrical, but this is unlikely. Therefore an attempt was made to find a symmetrical curve from the radio data alone, so that existing theory could be applied.

A number of trials were made to derive the best values for the systemic velocity, the position of the centre, and the orientation of the major axis; the first two were found from symmetry considerations, the last by maximizing the velocity variation. The analysis gave 170° for the position angle of the major axis, which is in satisfactory agreement with the optical value. The projected direction of the axis of rotation, the minor axis of the elliptical projection, is in position angle 80° .

The radio centre of rotation, *C_r*, at 05 hr 20 min, -68.8° , is, however, appreciably displaced from the optical centre *C* at 05 hr 24 min, -69.8° . The disagreement between the two centres is presumably associated with the asymmetrical shape of the Large Cloud; the position which was taken as the optical centre *C* is the centre of the innermost spiral pattern of dark matter and the geometrical centre of the axial bar, whereas the centre of rotation must be related to the centre of mass of the whole system. For comparison, the centroid *C_H* of the projected hydrogen distribution was found to be at 05 hr 35 min, -68.5° , and the centroid *C'* of the surface distribution of the brighter stars ($M \leq 4.7$) at 05 hr 33 min, -67.7° . The positions of the various centres are shown in Plate 1.

The above-mentioned values for position angle and centre of rotation were found to give a symmetrical velocity curve (see Figs. 3 (*a*), 3 (*b*)), and this will be used in the subsequent discussion of the mass of the Cloud, with points from the north and south sides of the system taken together to give the best mean curve (Fig. 4).

This curve gives a maximum value for the median rotational velocity of the Large Cloud of $17.5 \text{ sec } i \text{ km/sec}$ at 3.2° from the centre (or 2.5 kpc , taking the distance as 45 kpc). A Keplerian branch is well defined, extending out to about three times the distance of the velocity maximum. The shape of the curve is less reliable, however, in the outer parts where the lower intensity of the radiation leads to a greater scatter in the derived velocities.

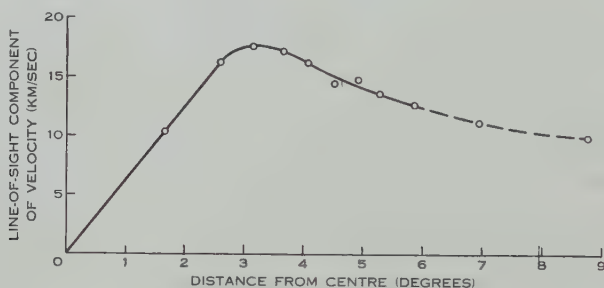


Fig. 4.—Mean rotational curve for the Large Cloud, obtained by taking the mean curve of Figure 3 (*b*) as symmetrical (based on median radial velocities). The plotted points are overlapping 15-point means.

The conversion from the line-of-sight component, $v_r \cos i$, to the actual rotational velocity, v_r , introduces a large uncertainty. The tilt angle of the Large Cloud is in the range where the secant varies rapidly with angle, so that a given proportional uncertainty in the tilt produces amplified effects in the rotational velocity and in the derived mass, which are proportional to $\sec i$ and $\sec^2 i$ respectively. The optically-derived value of the tilt, 65° , leads to a rotational velocity of 41 km/sec ; the velocity range corresponding to the $\pm 5^\circ$ probable error in tilt is $35\text{--}51 \text{ km/sec}$.

The radio rotational analysis can in principle provide an alternative method of deriving the tilt, but, as in the optical case, the method is very insensitive to variations of tilt when the system under investigation is nearly face-on. The radio results suggest that the tilt may be somewhat larger than 65° , but the present data are not sufficiently good to provide a satisfactory value.

The radio evidence for rotation gives support to the optically derived picture mentioned above, in which the Large Cloud is a flattened and tilted system showing extensive spiral structure. The sense of the tilt cannot be deduced from the optical evidence mentioned above, but an analysis of observations of Cepheid magnitudes (de Vaucouleurs 1954*a*, 1955*a*) suggests that the near side of the system is its north-following side. The sense of rotation derived on this assumption implies that the main spiral arm is trailing, as had already been suggested by the few optical velocities previously available.

(b) *Small Cloud*

The Small Cloud is tilted at a smaller angle to the line of sight (30°), so that the uncertainty in the tilt presents less difficulty in this case. The shape of the Small Cloud, however, has the complication that, in addition to the compact main body of the Cloud which dominates optical observations, there is a large prominence, or wing, extending towards the Large Cloud. This prominence is a striking feature of the 21 cm observations, although optically it is much narrower and somewhat shorter.

As with the Large Cloud, the median radial velocities for the Small Cloud were first analysed with respect to the optical centre, and then a new centre was derived from the radio data alone. The best rotational curve is obtained

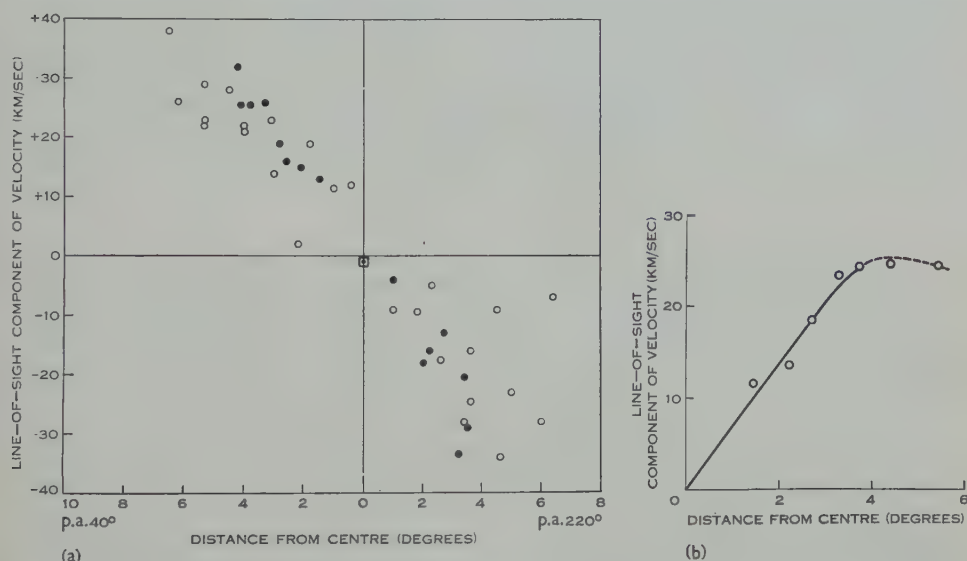


Fig. 5.—Rotational curves for the Small Cloud with respect to the derived *radio* centre of rotation, C_r , 01 hr 10 min, $-73\frac{1}{4}^\circ$, position angle of major axis $= 40^\circ$ (axis of rotation $= 130^\circ$) using *median* radial velocities. Systemic velocity, $v_s = -17$ km/sec.

- (a) Individual points; \bullet major axis $\pm 20^\circ$, \circ sectors 20° – 40° from major axis, \square mean of points along minor axis.
 (b) Mean curve; overlapping 10-point means.

with the centre C_r at 01 hr 10 min, $-73\frac{1}{4}^\circ$, and the major axis at position angle 40° (axis of rotation 130°), an orientation which agrees well with the optical value. A plot using these values, and covering a sector $\pm 40^\circ$ about the major axis, is given in Figure 5 (a), and the mean curve in Figure 5 (b). In this case, the observations were only just sufficient to go beyond the point of maximum velocity, but the curve is consistent with the assumption of rotation. The maximum rotational velocity for the Small Cloud is $26 \text{ sec } 30^\circ$, i.e. 30 km/sec, at a distance of 4.5° (3.5 kpc) from the centre.

The centre of rotation derived above is well displaced from the optical centre C , 00 hr 51 min, -73.1° , in the direction of the prominence, but the fact that it is displaced towards the centroid of the neutral hydrogen distribution,

C_H , 01 hr 20 min, -72.5° , supports the interpretation that it indicates the real centre of mass. For comparison, the centroid of the bright stars, C' , is at 01 hr 00 min, -72.5° (see Plate 2).

(c) *Reliability of Data*

An assessment of the internal consistency of the data can now be obtained from the deviations of the individual observed velocities from the mean rotational curve. The root-mean-square deviation is found to be 4.3 km/sec for the Large Cloud, and 6.7 km/sec for the Small Cloud. These values can be taken as upper limits for the relative observational errors, since they also include effects of systematic internal motions other than rotation.

The external consistency will be considered in Section IV, through a comparison with optical data.

Each observation refers to the region seen by an aerial whose half-power beam width is 1.5° . The observed velocity curves have therefore been smoothed to some extent by the aerial pattern, thereby lowering the velocity maximum. This effect will be corrected in a later paper which will discuss the masses of the Clouds.

IV. COMPARISON OF RADIO AND OPTICAL VELOCITIES

An upper limit for the external and absolute errors in the radio velocities can be obtained by comparing them where possible with available optical velocities (Wilson 1917). The agreement to be expected is limited by the fact that each optical velocity refers to a particular object, whereas the corresponding radio velocity refers to a mean over a large region. Further, as will be shown below, there are systematic differences between the optical velocities and the radio median radial velocities due to differences in effective depths in the Clouds. These systematic differences are additional to the effects of small-scale turbulence in the medium and the observational errors, so that the comparison sets only an upper limit to the combined radio and optical uncertainties. The r.m.s. value of the difference between corresponding radio and optical velocities is found to be 14 km/sec.

From a detailed examination (de Vaucouleurs 1955c) of the individual measurements of the optical spectra of Cloud nebulosities (Campbell and Moore 1918) the internal mean error of a velocity based on one plate of unit weight has been estimated at about 11 km/sec. Most optical velocities are based on three or four plates with a total weight between 2 and 3, hence have a mean error of the order of 8 km/sec, or 10 km/sec if the velocity dispersion amongst the objects themselves is included. Since the radio velocities have a mean error of the order of 5 km/sec (Section III (c)), one would expect an r.m.s. value of the order of 9 km/sec (11 km/sec including the cosmic dispersion) for the difference between radio and optical velocities. The fact that a slightly larger value (14 km/sec) is observed suggests that systematic differences might be present, which call for a closer scrutiny.

The mean difference between the 17 pairs of optical and radio velocities is +2 km/sec, which is a little less than the estimated probable error of such a mean. Hence the Doppler shifts at the two wavelengths are equal, within the

limits set by the observational uncertainty and the fact that different quantities are being measured in the two types of observation. This result is of philosophical interest because this is the first occasion on which it has been possible to compare directly Doppler shifts at radio and optical wavelengths, and also by inference the velocities of radio and light waves, outside the Earth-Moon system.

It is important to compare in greater detail the velocities measured by radio and optical methods, because along any particular line of sight the former refer to a weighted mean through the whole thickness of the system while the latter refer to a discrete sample of test objects presumably located in or close to the equatorial plane. This location is suggested by comparison with other spiral galaxies in which the stars which excite the luminescence of the emission nebulosities form the most flattened subsystem, and is indicated more definitely by the observed distribution of supergiant stars and emission nebulosities in the edgewise system of the Magellanic type NGC 55 (Plate 3). It is also strongly supported by the small velocity dispersion in the z -component (de Vaucouleurs 1955*c*).

TABLE 2
OPTICAL MEAN POINTS

Group No.	No. of Nebulae	Mean Position (1900)		Mean Velocity of Nebulae	Corresponding Radio Median Radial Velocity
		α (hr min)	δ		
1	4	04 53	$-69^{\circ} 24'$	+268	+266
2	3	04 55	$-66^{\circ} 44'$	+299	+281
3	4	05 27	$-68^{\circ} 04'$	+287	+281
4	6	05 41	$-69^{\circ} 38'$	+263	+272

Optical data are not available for a comparison with radio velocities in the Small Cloud, but in the Large Cloud the 17 emission nebulosities observed by Wilson (1917) can be used. The mean positions and the velocities for the four natural groups which they form are given in Table 2, together with the corresponding radio velocities. These optical values have been reduced, and corrected for the tilt of the Cloud, following the same procedure as for the radio data (Section III); they lead to the rotational velocities indicated by the squares in Figure 6.

We notice first that the optical points lie almost exactly on a straight line, which passes through the centre of symmetry of the radio curve. The most striking feature of Figure 6, however, is that the slope of the optical line is about twice that of the straight part of the radio median curve, indicating that the emission nebulosities are rotating more rapidly than the bulk of the gas. We infer that the emission nebulosities which are part of a more flattened subsystem share also in its greater spin, compared with the neutral hydrogen atmosphere.

Actually many of the 21 cm profiles for the Large Cloud are asymmetrical (Fig. 7). The radial velocity of the peak is up to 15 km/sec higher than the median radial velocity on the north side, and up to 15 km/sec lower on the south.

This velocity difference is about the same as that between the rotation curve of the emission nebulosities and that derived from the median velocities of the gas (see Fig. 6). We conclude that the gas near the equatorial plane of the Cloud is rotating more rapidly, the speed of rotation decreasing on going outwards from that plane. Along any line of sight through the Cloud the rotational velocity progressively increases to a maximum, and then decreases again. Since the density of gas presumably increases towards the equatorial plane, the peak

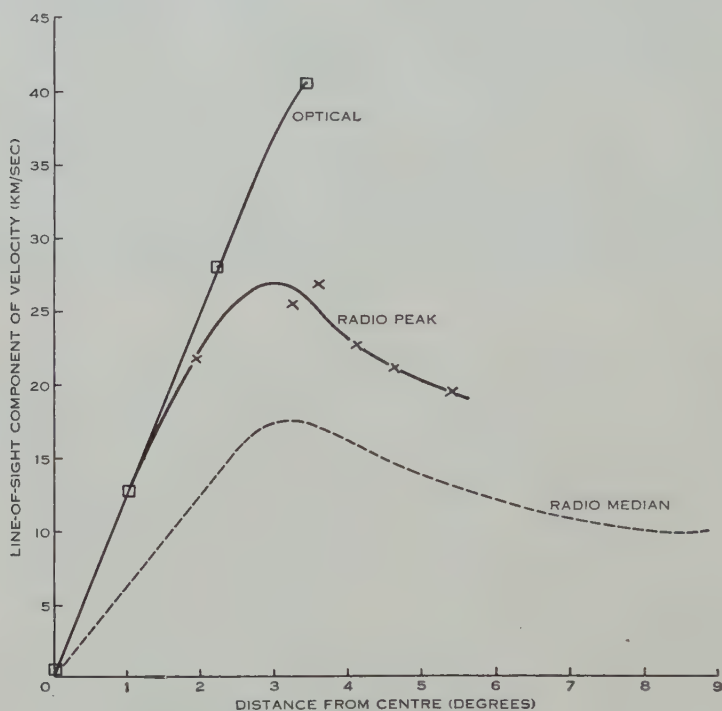


Fig. 6.—Comparison between rotation curves :

- Wilson's mean optical velocities of four groups of emission nebulosities (p.e. of mean points=4–5 km/sec),
- ×— radio peak velocities (p.e. of mean points=2–3 km/sec),
- radio median velocities (from Fig. 3 (b)).

of the observed profile should refer closely to the equatorial velocity, and will differ from the median velocity, which refers to some mean height $|z|$ above or below the equatorial plane, dependent on the z -distribution of densities and velocities.

The two curves are compared in Figure 6, indicating that the peak velocities are greater than the median velocities throughout; within 2° from the centre they are consistent with the optical velocities, which may then be used to deduce the probable course of the equatorial velocity curve near its origin. The greater optical velocities near $r=3^\circ$ may mean that the peak radio velocities do not refer to $z=0$ as do the optical values or that the small sample of nebulosities is affected by local departures. The difference, however, is partly due to the

smoothing of the radio-curve by the finite aerial beam and does not greatly exceed the combined uncertainty of both sets of data.

The line profiles in the weak outer parts are not yet good enough for a reliable determination of the peak velocities, so that only a small section of the equatorial rotation curve, near the velocity maximum, can be obtained at present.

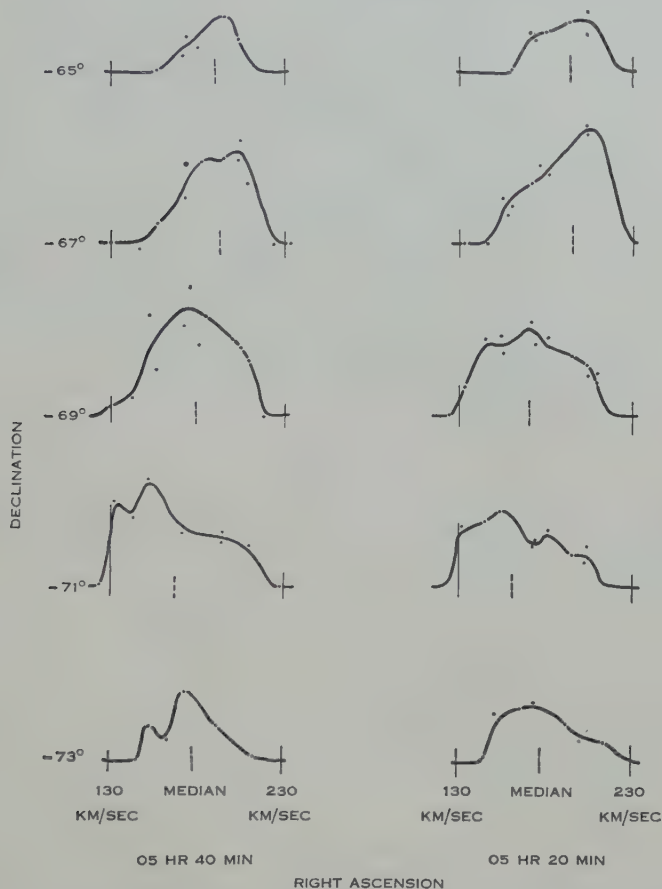


Fig. 7.—Examples of the Large Cloud profiles, showing the type of asymmetry observed. Values for the “peak velocity” could be identified in about 40 per cent. of the profiles which were obtained.

V. OTHER MOTIONS OF THE CLOUDS

The systematic radial velocity of the hydrogen of each Cloud can be derived by taking all the observed median radial velocities and weighting each one according to its associated brightness. The values obtained are $+280$ and $+161$ km/sec for the Large and Small Cloud respectively, with an estimated probable error of ± 1 km/sec. (For comparison, the median radial velocities at the centres of rotation are $+276$ and $+160$ km/sec respectively, and the mean of Wilson’s optical velocities $+276$ and $+168$ km/sec.*) These are the velocities

* Velocity of one nebulosity only in the case of the Small Cloud.

relative to the Sun ; after removal of the solar motion and a galactic rotation velocity of v_G , the systemic velocities become $(+280-0.90 v_G)$ and $(+161-0.65 v_G)$ respectively ; for $v_G=270$ km/sec, the values are $+37$ and -14 km/sec, for $v_G=220$ km/sec, $+82$ and $+18$ km/sec.

When more complete observational data become available, it may be possible to derive the actual space motion of each Cloud, and not merely its radial component. As each Cloud subtends a substantial angle, the component of a translational motion should vary across the Cloud (Hertzsprung 1920 ; Wilson 1944). The present data are not good enough for the rotational and translational motions to be separated, but this may be possible later.

The observed velocity dispersion (the half half-width of the line) in the Large Cloud has a mean value of about 25 km/sec, with slightly higher values in the central regions and slightly lower values in the outer parts. This velocity spread must be due in part to systematic mass motions in the equatorial plane, as shown by the irregularity of many of the profiles, and also to the variation of rotational speed with depth in the Cloud. There is also a small broadening due to the finite receiver bandwidth. Hence the true "random" velocities must be somewhat smaller than 25 km/sec, but the various effects cannot be separated on present evidence.

VI. CONCLUSIONS

The study of the 21 cm velocity distribution over the Large Cloud has shown that the Cloud is rotating, and the distribution for the Small Cloud is also consistent with a rotational interpretation. In each case the orientation of the derived axis of rotation agrees well with the minor axis indicated by the ellipticity of the outer optical isophotes. The radio and optical evidence taken together lead to a picture in which the Clouds are flattened systems inclined to the line of sight, and contain extensive spiral structure ; each Cloud bears some resemblance to a barred spiral, but its structure is asymmetrical and the centre of rotation appears to be displaced from the centre of the axial bar. The neutral hydrogen subsystem is less flattened and less concentrated radially than the other constituents and thus may be regarded as a gaseous envelope.

The most extensive rotation curve for the Large Cloud has been derived from the median radial velocity for each profile. A comparison with a smaller number of optical and radio peak velocities has demonstrated the variation of rotational speed with height above and below the equatorial plane and shown that the median velocities refer to an unknown mean level which is displaced from that plane. Hence a rotational determination of the mass of the Large Cloud, which will be considered in a later paper, should be based on a velocity curve derived from the peaks of the line profiles, i.e. the equatorial velocity curve, rather than on the median velocities. The peak velocity information is not yet sufficiently extensive to yield an independent rotation curve, but the median velocities can be used to indicate the overall shape of the curve.

Besides the rotation, the radio velocities have given the radial component of the space motion of each Cloud and have shown that the internal random velocities are about 20 km/sec.

VII. REFERENCES

- BAADE, W., and MAYALL, N. U. (1951).—"Problems of Cosmical Aerodynamics." pp. 165-84.
(U.S. Central Air Documents Office: Dayton, Ohio.)
- BABCOCK, H. W. (1939).—*Bull. Lick Obs.* **19**: 41-51.
- BUSCOMBE, W., GASCOIGNE, S. C. B., and DE VAUCOULEURS, G. (1954).—*Aust. J. Sci.* **17** (3)
Suppl.: 1-35.
- CAMPBELL, W. W., and MOORE, J. H. (1918).—*Publ. Lick Obs.* **13**: 77-183.
- HERTZSPRUNG, E. (1920).—*Mon. Not. R. Astr. Soc.* **80**: 782-4.
- KERR, F. J., and HINDMAN, J. V. (1953).—*Astr. J.* **58**: 218-9.
- KERR, F. J., HINDMAN, J. V., and ROBINSON, B. J. (1954).—*Aust. J. Phys.* **7**: 297-314.
- LUYTEN, W. J. (1928).—*Proc. Nat. Acad. Sci., Wash.* **14**: 241-5.
- MAYALL, N. U. (1946).—*Astrophys. J.* **104**: 290-323.
- MAYALL, N. U. (1950).—*Publ. Obs. Univ. Mich.* **10**: 19-24.
- MAYALL, N. U., and ALLER, L. H. (1942).—*Astrophys. J.* **95**: 5-23.
- MILLS, B. Y. (1955).—*Aust. J. Phys.* **8**: 368-89.
- DE VAUCOULEURS, G. (1954a).—*Observatory* **74**: 23-31.
- DE VAUCOULEURS, G. (1954b).—*Observatory* **74**: 158-64.
- DE VAUCOULEURS, G. (1955a).—*Astr. J.* **60**: 126-40.
- DE VAUCOULEURS, G. (1955b).—*Astr. J.* **60**: (in press).
- DE VAUCOULEURS, G. (1955c).—Velocity dispersion in the Large Magellanic Cloud from the Lick
velocities of emission nebulosities. *Publ. Astr. Soc. Pacif.* (in press).
- WILSON, R. E. (1917).—*Publ. Lick Obs.* **13**: 187-90.
- WILSON, R. E. (1944).—*Publ. Astr. Soc. Pacif.* **56**: 102-6.

APPENDIX I

Determination of Rotational Velocity from Observed Radial Velocity, allowing for Tilt

Let ABC (Fig. 8) be the elliptical projection on the celestial sphere of a circle of radius r in the equatorial plane of the system, which is inclined at an angle i to the line of sight. Let P_0 be a point on this ellipse at azimuth angle θ_0 and apparent distance r_0 from the centre. Then the observed radial velocity

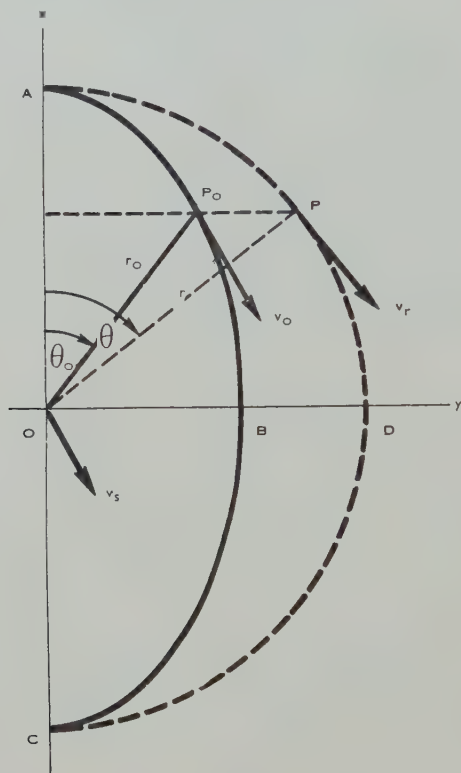


Fig. 8.—Diagram illustrating the method of correcting the observed radial velocities for the tilt of the system.

v_0 at P_0 is related to the rotational velocity v_r of the point P in the corresponding circle by the expression

$$v_0 - v_s = v_r \cos \theta \cos i,$$

where v_s is the radial velocity of the system as a whole (systemic velocity), and θ is obtained from the observed azimuth angle θ_0 by

$$\tan \theta = \tan \theta_0 \operatorname{cosec} i.$$

In this way a value of v_r can be derived from each observed v_0 . The corresponding distance from the centre r is related to the apparent distance r_0 by

$$r = r_0 \sec \theta / \sec \theta_0.$$

MOVEMENT OF SPORADIC E IONIZATION

By J. A. HARVEY*

[*Manuscript received July 18, 1955*]

Summary

The horizontal movement of patches of day-time sporadic E ionization has been observed by using a system of spaced pulse transmitters and a central recorder. The directions of movement are mainly towards the north and west, with speeds mainly between 40 and 80 m/sec. These velocities differ in direction from the F region disturbances recorded at the same time and have only about half the speed.

I. INTRODUCTION

During the last few years, evidence for movement of patches of sporadic E ionization (hereafter called E_s patches) has come from several sources. Ferrell (1948) and Gerson (1950) demonstrated the movement of large patches of E_s by using the observations of amateur radio operators. Bramley (1953) tracked small patches by using direction finders and Findlay (1953) observed the movement of small patches by using a phase path method.

Gerson calculated that the E_s patch movements had speeds of about 50–60 m/sec in various directions. Bramley, from only a small number of instances, found speeds of between 40 and 60 m/sec in various different directions, while Findlay's calculations, for speeds only, gave results of the same order.

Using a system of spaced pulse transmitters and a common receiver, records of virtual height against time ($h't$) have been made during the day-time at Sydney for the purpose of studying types of horizontal movements occurring in the ionosphere. Some results of these experiments, for the F region, have been described by Munro (1950). For periods of minutes to hours, E region echoes from heights of about 110–120 km frequently appear on these records. The times of occurrence of such echoes, as observed for the three transmissions, differ by some minutes.

Except in about 5 per cent. of the cases considered in this paper, the echoes show a constant height for each E_s patch. The exceptions show an apparent fall in height after the first appearance of the echo and a corresponding rise as the echo disappears. During these changes in height, which take about 10–15 min, the echoes are weak. Assuming that these E_s patches have a lenticular shape, the apparent change in height shown by the echo would plausibly indicate reflection from the edge of an approaching (or receding) E_s patch; the rate of change of height would give a measure of the horizontal speed of the patch. However, due to difficulties of measurement only very approximate speeds can be obtained in this way; these range between 50 and 100 m/sec.

* Radio Research Board, C.S.I.R.O., Electrical Engineering Department, University of Sydney.

More accurate velocities of movement of the E_s patches can be determined from the time differences of appearance (and disappearance) of the echoes from the three transmitters. Velocities so obtained for the period from November 1950 to October 1952 are discussed in this paper.

Also present on the records are E region echoes which appear simultaneously from all three transmitters. They show virtual heights of about 100 km and last only a few seconds. These are almost certainly due to transient ionization produced by meteors; they will not be considered further in this paper.

II. OBSERVING SYSTEM

Figure 1 shows the relative positions of the three pulse transmitters, the recorder, and the corresponding reflection points, assuming a plane and horizontal ionosphere. Transmitter T_1 and the recorder are located at the University of Sydney, T_2 at Camden, and T_3 at Blaxland. The transmitters operate on a

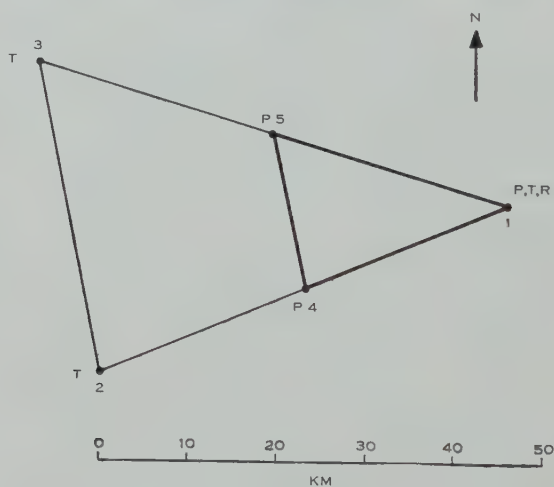
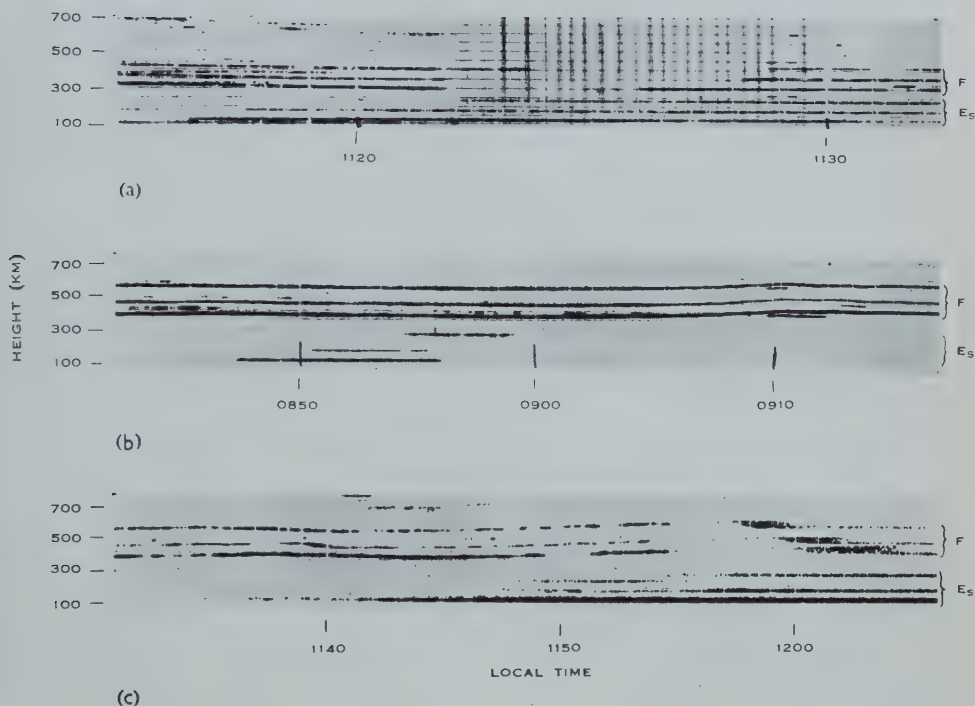


Fig. 1. — Plan of observing system. R , recorder; T , transmitter; P , projection of reflection point. Points 1, 4, and 5 form the observing triangle.

frequency of 5.8 Mc/s with a pulse duration of 30 μ sec, a pulse repetition frequency of 50/sec, and a peak power output of 1 kW. Horizontal half-wave aeriels are used and the effective power radiated vertically is approximately the same for all three. They irradiate the sky directly above them uniformly within a cone whose aperture is 130° in the vertical plane perpendicular to the aerial and 90° in the vertical plane containing the aerial. The remote transmitters are slave-triggered by the ground pulse from the master transmitter situated at point 1. Triggering delays of about 400 and 1000 μ sec are introduced so that the echoes due to the three transmitters are suitably spaced on the record. For timing purposes, and to permit identification of the echoes due to each transmitter, T_1 is interrupted every minute for a few seconds; T_2 every minute (but not synchronously with T_1); and T_3 every half minute.

MOVEMENT OF SPORADIC E IONIZATION



Examples of echoes from E_s patches. From bottom to top of each example the order of the echoes from the three transmitters is T_1 , T_2 , and T_3 . The height scale applies only to T_1 , the pulses from the other transmitters being delayed so as to separate their echoes on the record. (a) August 20, 1951: a set of clear E_s echoes. The F echoes are cut off for a few minutes. (b) December 25, 1951: a set of clear echoes from a small patch. (c) December 31, 1951: a set of slightly diffuse E_s echoes. These show a slight fall in height particularly for the echo from T_1 .

A typical F disturbance has occurred near the right-hand end of the F traces.

For reception a half-wave aerial is used, oriented approximately at right angles to the mean direction of the remote transmitters so as to equalize their signals. The echoes, together with height calibration marks, are displayed on a cathode-ray tube intensity modulated. The display is photographed on 35-mm film moving slowly at right angles to the time base, so producing an $h't$ record of the three sets of echoes.

III. CHARACTERISTICS OF THE E_s PATCHES

The echoes from the E_s patches appear on the records either as clear traces or as diffuse traces. The E_s echoes producing clear traces show a spread of 10 km equivalent height, this being the sharpest trace obtainable with the recording equipment used. They appear and disappear, sometimes very rapidly, with no change in height. The F echoes are frequently cut off, sometimes very rapidly and in a number of cases for nearly the full duration of the E_s echoes. Fading is slow, with periods ranging from a few seconds to one or two minutes. In Plate 1, (a) and (b) illustrate two examples with these properties. Diffuse E_s echoes, which always appear and disappear gradually, are spread over as much as 50 km equivalent height. The F echoes are often blanketed when this occurs, the onset of blanketing being gradual. Fading is always fast, with periods of less than a second, different parts of the spread echo fading at different rates. Some diffuse echoes show a fall in height during the first 15 min after appearance and a similar rise before disappearing. This change in height is usually about 20–30 km but may be as much as 50 km; the echoes are weak at such times. In Plate 1, (c) shows a slightly diffuse echo with a small change in height taking place during the 8 min following its first appearance.

These properties are consistent with the assumption that E_s patches are sheets (or clouds) of moving ionized particles. The occurrence of clear traces showing no height changes suggests that the E_s patches sometimes have surfaces that are flat and smooth enough to give only specular reflections. The diffuse traces suggest that at other times the surfaces are sufficiently rough to give scattered, rather than specular, reflections. The initial fall in height, together with the accompanying small amplitude of some of the diffuse echoes, also suggests that these patches have sufficient thickness to give weak edge scattering when as far as 120 km horizontally from the recorder. (The low amplitude of these echoes is not due to the aerial radiation patterns—see Section II.)

The heights of the E_s patches lie mainly between 110 and 120 km with limits of 100–140 km; $h'f$ records show that they lie just above the normal E layer, which is usually at a height of 100–110 km.

IV. ANALYSIS OF RECORDS

In order to calculate the velocities of the E_s patches from the differences in time of appearance (or disappearance) of the echoes, it is necessary to assume a definite shape for the E_s patch. The simplest assumption is to assume that the E_s patch has leading and trailing edges which are straight and parallel, that it is moving with constant velocity, and that it is fixed in size. This enables the use of a simple graphical method for calculating the velocities of movement.

The validity of this assumption was tested by comparing, for each E_s patch, the velocities calculated in three different ways, namely, from the times of appearance of the three echoes, from the times of their disappearance, and from the times of occurrence of the mid points of the three echo traces. If this assumption is valid, these three velocities should be identical.

Figure 2 is a diagrammatic representation of a typical record of the echoes from a small patch of E_s similar to the example shown in Plate 1 (b). Lines 1, 2, and 3 represent the echoes of pulses from the three transmitters reflected from the E_s patch, while the t 's are the various times of occurrence indicated

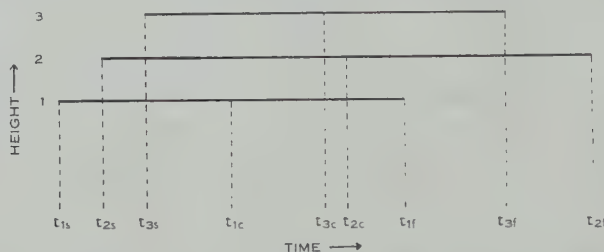


Fig. 2.—This diagram represents the echoes from an E_s patch, similar to the examples in Plate 1 and shows the various times referred to in the text.

by the suffixes s for the start, f for the end, and c for the mid point of each echo trace. The starting time differences are then defined as :

$$T_{2s} = t_{2s} - t_{1s}, \quad T_{3s} = t_{3s} - t_{1s},$$

and similarly the ending time differences. The centre time differences are then determined from the relation :

$$T_{2c} = (T_{2s} + T_{2f})/2, \quad T_{3c} = (T_{3s} + T_{3f})/2.$$

From these time differences, the corresponding velocities of movement can be calculated. As the patches are assumed horizontal and the reflection specular, negligible error is incurred by assuming that the reflection points are directly above the half-way points between the remote transmitters and the recorder. Also, as the angle of incidence of the rays from the remote transmitters is about 15° , the possible error due to obliquity is also negligible.

The time of the start and the finish of each echo trace was noted. For all patches where the start and finish times for all three echo traces could be read to an accuracy of 1 min or better, the time differences were calculated. Many of the patches whose echoes showed diffuseness, including all those which showed change of height, and also some which showed no diffuseness, were thus rejected. The corresponding directions and speeds were then determined from the time differences for the remaining patches.

It was found that the directions and speeds, as calculated from the beginning, centre, and ending of the echo traces from an E_s patch, were usually different. The directions differed by anything from 0° to 180° while the ratio of the speeds

varied between 1 and 12 or occasionally more. Figure 3 shows a histogram of the distribution of the direction differences for the period from November 1951 to February 1952. For approximately 60 per cent. of the patches, it shows that the start and finish directions differ by 60° or less but for the remainder the direction differences are fairly evenly distributed up to 180° . It appears then that the simple assumptions about shape and movement made above do not always apply. Possible causes of the differences are that the E_s patches have curved edges, are moving with variable speed, and changing in size and shape. However, it is not possible to determine which of these causes is responsible or to make any estimate of the shape of the patches from observations taken at only three points.

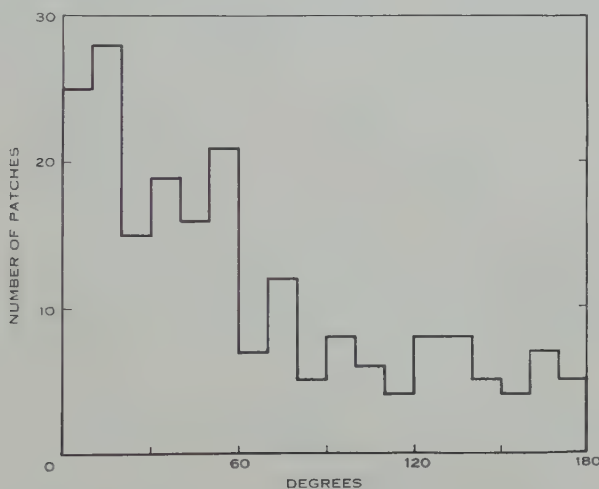


Fig. 3.—Distribution of the difference between the beginning and ending directions of each E_s patch, from November 1951 to February 1952.

It seemed more reasonable to take the velocities of the centres of the patches rather than those of the edges. This reduced errors due to possible differences in output from the three transmitters and any residual effects of obliquity, all other equipment effects being common to the three signals. The advantage of this method was emphasized by a numerical study of the velocities deduced for various regular-shaped patches crossing the observing network, which revealed that the centre velocity was consistently more reliable than the velocities of the edges. In particular, for triangular and circular patches moving at constant velocity, the centre velocities are correct, and, for constant rate of change of size as well, almost correct even though the beginning and ending velocities may be widely different, particularly in the direction component.

In order to establish a criterion for selecting the velocities thus determined so that E_s patches with predominantly horizontal movement rather than change in size or shape would be used, the numerical calculations of velocities for the assumed regular-shaped patches and the observed velocities of the E_s patches were compared. From this it was decided to select those E_s patches whose

beginning and ending velocities differed in direction by less than 80° and with speed ratios less than 5 to 1. This eliminated about 35 per cent. of the patches. Figure 4 shows the difference between using the beginning and ending velocities for all the patches, and the velocities of the centres of the selected patches for June 1952. It will be seen that the scatter of points is considerably reduced by using the selected velocities.

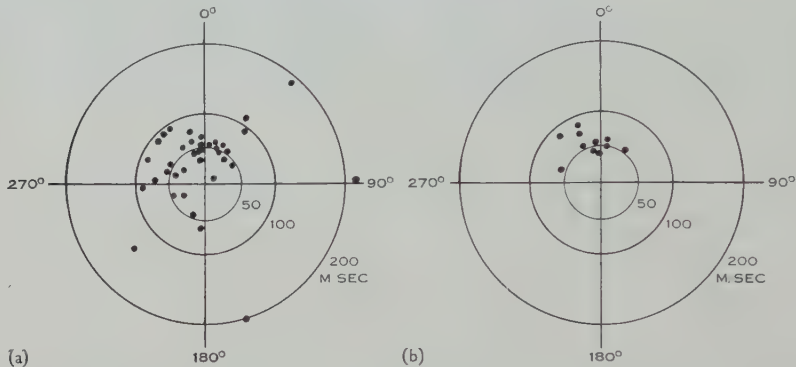


Fig. 4.—This diagram shows the reduction of scatter of points obtained by using the velocities of the centres of the selected E_s patches instead of the velocities of all the patches recorded in June 1952. (a) All E_s patches using beginning and ending velocities; (b) selected E_s patches using centre velocities.

V. RESULTS

The recording period was from 0600 to 1800 hr daily, Eastern Australian time, and these records were examined for E_s times of appearance and disappearance for the period from November 1950 to October 1952. At 5.8 Mc/s during this period, the E_s was most prevalent during the summer months of November-February. A small subsidiary maximum occurred during the winter of 1951 (May-August) but not in 1952; during July and August there was much less E_s than during the same months of the previous year. E_s was almost completely absent in March and April of 1951, and in September and October of both years. The velocities of the E_s patches were then determined and selected as described in Section IV. About 65 per cent. of the patches whose velocities could be determined were accepted as suitable for investigation. They are discussed below, using centre velocities only.

Figure 5 shows the distribution of occurrence of the E_s patches during the day from November 1951 to February 1952. The E_s was most prevalent during the late morning. A similar distribution was found for the previous summer. For the winter, the E_s occurred mainly between 1200 and 1500 hr.

It was not possible to get much information on the size and shape of the patches using only three observing points. However, from a study of the differences in directions and speeds as determined from the beginning and ending of the echo traces, it seemed probable that the edges of the E_s patches were curved rather than straight, and that the patches usually changed in size or

shape while crossing the observing points. In size they must have been at least as wide as the observing triangle, which is of the order of 25 km, otherwise

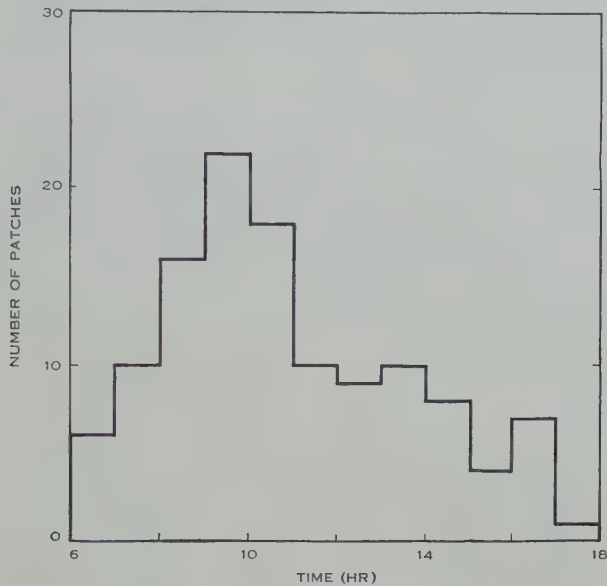


Fig. 5.—Distribution of the times of occurrence of the E_s patches, from November 1951 to February 1952.

specular reflections from all three transmitters would not occur. An approximate value for the length of a patch was obtained by taking the product of the speed

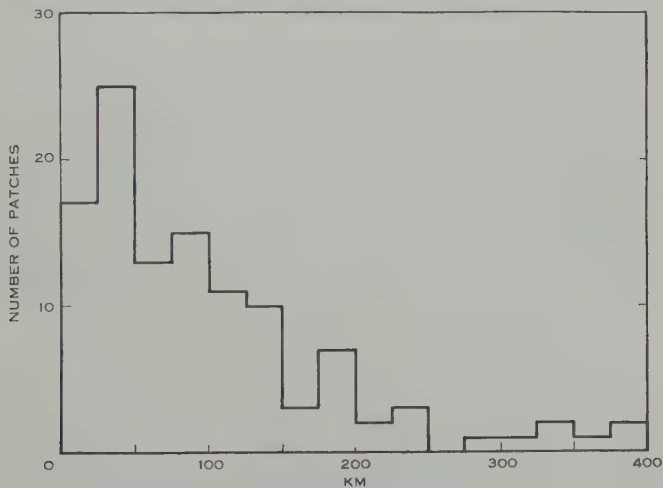


Fig. 6.—Distribution of the lengths of the E_s patches, from November 1951 to February 1952.

of its centre and the time taken for it to cross the recording point. Figure 6 shows the distribution of these lengths over the four summer months from November 1951 to February 1952.

This figure also shows that 13 per cent. of the patches have lengths less than 25 km. This means these patches are shorter than they are wide, the lengths of a few being as short as 6 km. Of those longer than 25 km, the length/width ratio is indeterminate, hence it is not possible to compare the relative proportions of short and long patches.

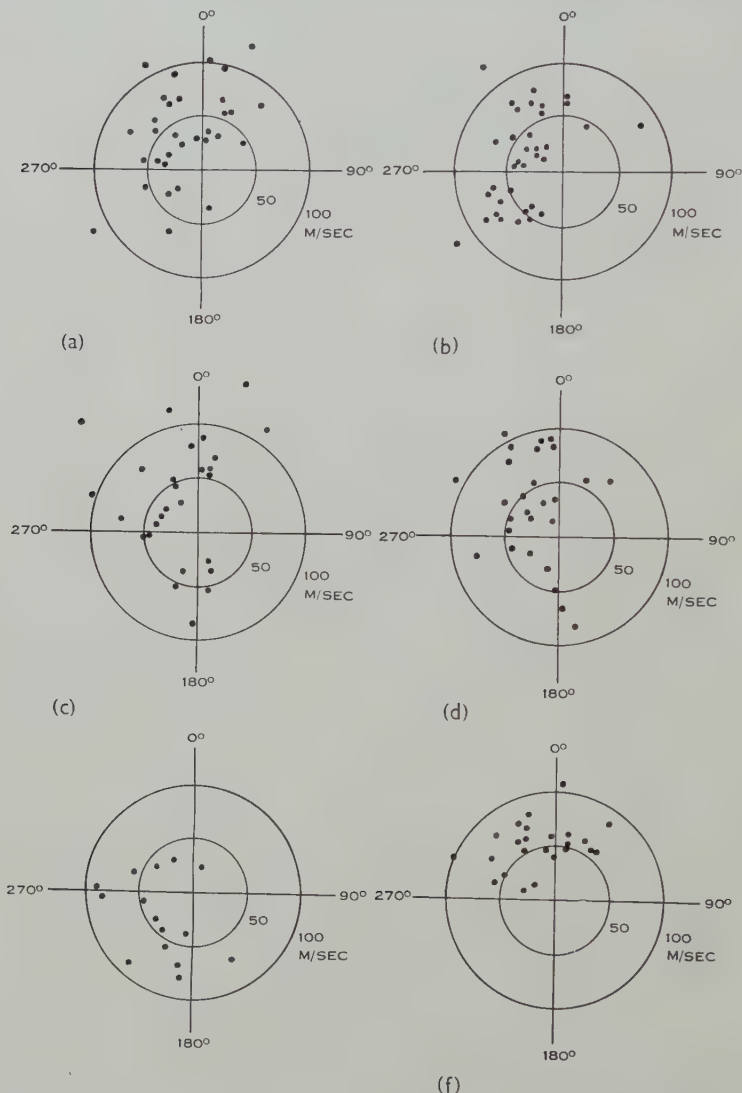


Fig. 7.—Velocities of the E_s patches. (a) November 1951; (b) December; (c) January 1952; (d) February; (e) March and April; (f) May-August.

(a) Velocities

Figures 7 (a)–7 (d) show polar plots of the velocities of the E_s patches for the four summer months from November 1951 to February 1952. The directions

are the vector directions of movement measured in degrees east of north and the speeds are measured in m/sec. The directions lay mainly within the sector between 270° and 030° . In December, another group appeared around a mean direction of 240° ; in January, there was a small group around 180° . The speeds lay mainly between 20 and 80 m/sec. There was no movement of patches in directions lying between 60° and 150° .

In winter, the directions of movement were similar to those in summer. Figure 7 (f) for May-August 1952 shows that the directions lay almost completely in the sector between 285° and 45° , the scatter being less than that for the summer. For the equinoctial months of March and April 1952, Figure 7 (e) shows that the directions were grouped around 225° but that the scatter was considerable.

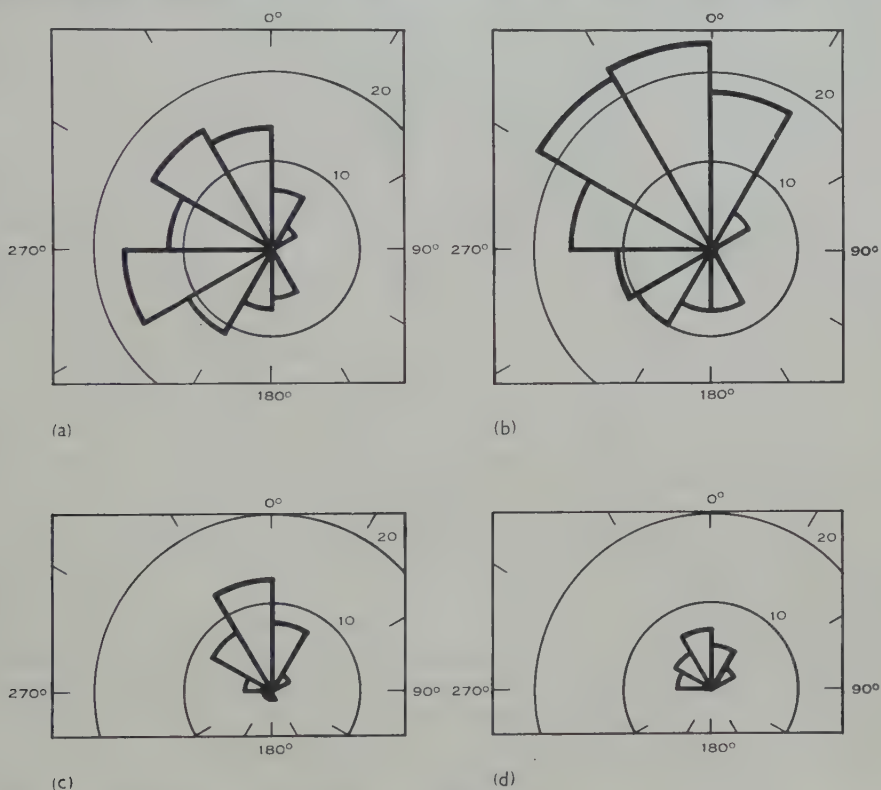


Fig. 8.—Distribution of directions of the E_s patches. (a) Summer 1950–51; (b) summer 1951–52; (c) winter 1951; (d) winter 1952. Circles: number of patches.

(b) Comparison of 2 Years' Data

For summer, Figures 8 (a) and 8 (b) are histograms of the distribution of directions for 1950–51 and 1951–52. During the summer of 1950–51, the directions were more westerly than during the summer of 1951–52. Figures 8 (c) and 8 (d) show the distributions for the winters of 1951 and 1952. Both distributions are very similar. In all seasons, summer, equinox, and winter, there was no movement of patches in directions lying between 60° and 150° .

Figures 9 (a)–9 (d) show histograms of the distribution of speeds for the above periods. The bulk of the speeds fell in the range from 20 to 80 m/sec. In 1951, there was a larger proportion of higher speeds than in 1952.

(c) *Seasonal and Diurnal Variations*

There are not enough data available to reveal any definite seasonal or diurnal changes. The only indication of seasonal change appears to be at the March-April equinox of 1952. Here there is a swing of directions of about 90°

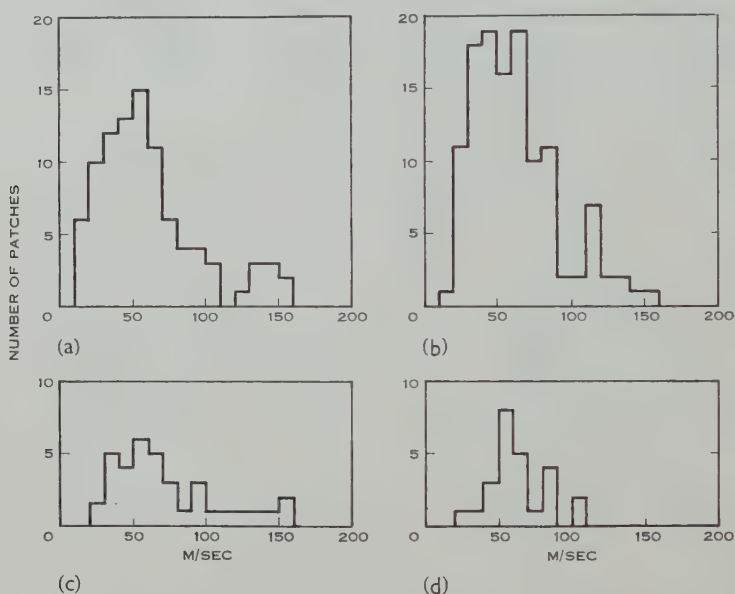


Fig. 9.—Distribution of speeds of the E_s patches. (a) Summer 1950–51; (b) summer 1951–52; (c) winter 1951; (d) winter 1952.

towards the south-west. For the same period of 1951 and the September-October equinox in both years, there was not enough E_s to give adequate statistics. As the E_s was recorded only in the day-time and occurred mainly over the limited period of 0800–1200 hr in summer, and 1200–1400 hr in winter, no diurnal variation could be found.

VI. DISCUSSION OF RESULTS

The movements of E_s patches take place in the E region just above the normal E height which is usually about 100–105 km. The speeds of movement are similar to those found using different methods and in other places, such as the fading of echoes by Mitra (1949), Salzberg and Greenstone (1951), and Phillips (1952); observations of small E_s patches using a phase path method by Findlay (1953); direction-finding of remote transmitters by Bramley (1953); and observation of meteor trails by Manning, Villard, and Peterson (1950) and by Elford and Robertson (1953). Briggs and Spencer (1954) have discussed the various methods of measuring horizontal movements, particularly the fading method using closely spaced receivers. For the E region, they find a most

probable speed of 80 m/sec; this is higher than our most frequent speed (60 m/sec) for the E_s patches.

Comparison in detail of the directions of movement of the E_s patches with those obtained by other methods is generally difficult owing to the limited amount of data for the E_s patches. The winds obtained by Elford and Robertson of Adelaide (which has a similar latitude to Sydney) for the months of November and December 1952, and between 0900 and 1200 hr, were mainly in the direction 90° . Briggs and Spencer quote similar directions for summer at Cambridge. This is almost directly opposite to the directions of the E_s patches, which are mainly between 270 and 360° . No patches were found moving in the directions between 60 and 150° . For winter, Briggs and Spencer quote directions which show a steady change through the day and which are 300 , 330 , and 030° , corresponding to 240 , 210 , and 150° for the southern hemisphere at 1200, 1300, and 1400 hr respectively. This compares approximately with our 340° for the E_s patches at the same time of day. There appears to be a westerly component of velocity in both cases; our north-south component is towards the north instead of the south.

In both seasons the directions of movement of the E_s patches are markedly different from those found by the echo-fading method, and by reflections from meteor trails. This suggests that there is a change of direction of movement with height as the E_s patches lie above the region (between 85 and 100 km) where meteor trails occur, and also above the region producing the fading of echoes. This latter cannot be higher than the E layer and may be lower.

Despite the differences in directions, the similarity of speeds suggests that there is a common cause for movements of E_s patches, fading patterns, and meteor trails. This cause could be winds. Martyn (1953) has discussed the drift of ionization due to winds in the presence of the Earth's magnetic field. At Sydney, the effect of the magnetic field on the motion of ionization is negligible up to a height of 120 km; hence the E_s patches move with the wind; above 120 km the effect of the magnetic field becomes more appreciable.

Clemmow, Johnson, and Weekes (1955) have discussed a steady state solution to the problem of the motion under the influence of a wind of a cylinder of ionization differing in density to the surrounding ionized medium and in the presence of a magnetic field. For the upper E region, a dense cylinder of ionization will move with a velocity different from that of the wind, but the difference is not large.

Both these effects seem insufficient to account for the wide differences between the directions of movement of the E_s patches, and of the fading patterns and the drift of meteor trails. Thus the cause of all the movements appears to be winds whose velocity changes with height.

VII. COMPARISON WITH F REGION DISTURBANCES

From the same records, the movements of F region disturbances as described by Munro (1950) are quite different. Table 1 shows the median values of direction and speed of both the E_s patches and the F disturbances for two summers and two winters.

The speeds of F disturbances are about twice those of the E_s patches. In summer the F directions are almost directly opposite to those of the E_s patches and lie mainly in the sector where there are no E_s directions. In winter the difference is only about 60° , the F directions having swung round about 90° to the north while the E_s directions have changed only a little.

TABLE I
VELOCITIES OF E_s PATCHES AND F DISTURBANCES
Median values

Season	E_s Patches		F Disturbances	
	Direction ($^\circ$ E. of N.)	Speed (m/sec)	Direction ($^\circ$ E. of N.)	Speed (m/sec)
Summer 1950-51 ..	280	58	125	120
Summer 1951-52 ..	315	57	120	125
Winter 1951	340	60	030	142
Winter 1952	330	60	025	145

VIII. CONCLUSIONS

The results show that there is definite movement of ionization or ionizing sources in the E region. The similarity of the speeds to those obtained both by the fading method and by observations of meteor trails, indicates a common cause which is probably winds. There is some seasonal change of these winds.

IX. ACKNOWLEDGMENTS

The author wishes to thank Dr. D. F. Martyn, F.R.S., Chief Scientific Officer, and Dr. G. H. Munro, both of the Radio Research Board, for their helpful discussions and advice. The facilities provided by the Department of Electrical Engineering of the University of Sydney are gratefully acknowledged.

X. REFERENCES

- BRAMLEY, E. N. (1953).—*Proc. Roy. Soc. A* **220** : 39.
 BRIGGS, B. H., and SPENCER, M. (1954).—*Rep. Progr. Phys.* **17** : 245.
 CLEMMOW, P. C., JOHNSON, M. A., and WEEKES, K. (1955).—"The Physics of the Ionosphere," p. 136. (Physical Society : London.)
 ELFORD, W. G., and ROBERTSON, D. S. (1953).—*J. Atmos. Terr. Phys.* **4** : 271.
 FERRELL, O. P. (1948).—*Proc. Inst. Radio Engrs., N.Y.* **36** : 879.
 FINDLAY, J. W. (1953).—*J. Atmos. Terr. Phys.* **3** : 73.
 GERSON, N. C. (1950).—*Nature* **166** : 316.
 MANNING, L. A., VILLARD, O. G., and PETERSON, A. M. (1950).—*Proc. Inst. Radio Engrs., N.Y.* **38** : 877.
 MARTYN, D. F. (1953).—*Phil. Trans. A* **246** : 306.
 MITRA, S. N. (1949).—*Proc. Instn. Elect. Engrs.* III **96** : 441.
 MUNRO, G. H. (1950).—*Proc. Roy. Soc. A* **202** : 208.
 PHILLIPS, G. J. (1952).—*J. Atmos. Terr. Phys.* **2** : 141.
 SALZBERG, C. D., and GREENSTONE, R. (1951).—*J. Geophys. Res.* **56** : 521.

SOME OBSERVATIONS OF WIND VELOCITY AUTOCORRELATIONS IN THE LOWEST LAYERS OF THE ATMOSPHERE

By R. J. TAYLOR*

[*Manuscript received May 16, 1955*]

Summary

From wind velocity observations in the lowest layers of the atmosphere, mean square velocity differences over a time interval at a point fixed in space are derived and their variation with the time interval is considered. The magnitude of these mean square differences is related to the rate of viscous dissipation of energy and to the shearing stress. On the average, fair agreement with predictions from the dimensional arguments of the theory of local isotropy is shown even though the results pertain to eddy sizes outside the inertial subrange as usually defined.

I. INTRODUCTION

The theory of local isotropy attributed to Kolmogoroff and first expounded in English by Batchelor (1947) is generally taken as applying only to the "viscous" and "inertial" subranges of eddy sizes and special interest attaches to the second of these, that is, to those eddies which are small enough to be statistically decoupled from the larger, directionally biased eddies and yet not so small that their motion is appreciably affected by viscous forces. The theory is not intended to apply to eddies larger than those in this subrange. Nevertheless autocorrelations with regularities of the type predicted by this theory have been observed in the atmosphere (e.g. Taylor 1952*a*; MacCready 1953) up to eddy sizes several times the height of observation where one would expect the influence of the underlying surface in imposing a directional bias to outweigh any inherent tendency of the motion towards isotropy.

It must be realized that the theory of local isotropy in its usual form is divisible into two parts: first, the two separate hypotheses of statistical decoupling and local isotropy proper and, secondly, the dimensional arguments based on them. It would appear that, while these hypotheses are a logical basis for the dimensional arguments, they may not be the only possible ones and that these arguments may be valid for other reasons even outside the domain where a Kolmogoroff "inertial subrange" can be expected to exist. The purpose of this paper is to present a series of results which, though showing wide scatter, in the mean exhibit this type of unexpected agreement with the dimensional arguments of the theory of local isotropy.

II. RESULTS

As part of the programme of field work of the C.S.I.R.O. Division of Meteorological Physics, a large number of photographically recorded galvanometer traces of the speed, azimuth, and inclination to the horizon of the wind has

* Division of Meteorological Physics, C.S.I.R.O., Melbourne.

been made. The techniques involved and other analyses of these records are described elsewhere. By means of a process of mechanical computation, it is possible to rewrite this information in the form of continuous traces of u , v , and w ; respectively the down-wind, cross-wind, and vertical components of velocity. Each of the records used here represents 5 min of recording time.

Autocorrelation coefficients for a number of different time lags have been calculated from the traces of u , v , and w for a total of nine 5-min runs and are exhibited in Table 1 together with relevant meteorological data. The time obtained by dividing the height of observation by the mean wind speed has also been tabulated as an indication of the time of passage of an eddy of down-wind dimension equal to the height of observation.

Since the original records were made with critically damped galvanometers having natural periods of approximately 2 sec, it might be suspected that false autocorrelations are thereby introduced into the record. This was investigated by considering the equation describing the response, $\theta(t)$, of a critically damped galvanometer of natural period $2\pi/\omega$ to a fluctuating input $u'(t)$, where θ and u' are to be considered as excursions from mean values. This equation is

$$\ddot{\theta} + 2\omega\dot{\theta} + \omega^2\theta = \omega^2 u'(t).$$

We define functions

$$R_u(\sigma) = \overline{u'(t)u'(t+\sigma)},$$

similar functions involving v and w ,
and

$$\rho(\sigma) = \overline{\theta(t)\theta(t+\sigma)},$$

where a bar denotes a mean taken over the whole record. These represent the time lag covariances existing in the original quantity and in the galvanometer trace respectively. It follows then that

$$\begin{aligned} R = \omega^{-4} & \left[\overline{\ddot{\theta}(t)\ddot{\theta}(t+\sigma)} + 2\omega\{\overline{\ddot{\theta}(t)\dot{\theta}(t+\sigma)} + \overline{\dot{\theta}(t)\ddot{\theta}(t+\sigma)}\} \right. \\ & + \omega^2\{\overline{\ddot{\theta}(t)\theta(t+\sigma)} + 4\overline{\dot{\theta}(t)\dot{\theta}(t+\sigma)} + \overline{\theta(t)\ddot{\theta}(t+\sigma)}\} \\ & \left. + 2\omega^3\{\overline{\dot{\theta}(t)\theta(t+\sigma)} + \overline{\theta(t)\dot{\theta}(t+\sigma)}\} + \omega^4\overline{\theta(t)\theta(t+\sigma)} \right]. \end{aligned}$$

We can now make use of a type of argument introduced by G. I. Taylor (1922) to investigate certain properties of a fluctuating quantity. Following him, we suppose that a value of ρ is calculated by reading off values of θ from the record at a large but finite number of points and evaluating the form for ρ given above. Then we can expect the value of ρ to remain unaltered if a different set of points, displaced in time by a constant small amount δt from the first set, is used instead, that is,

$$\begin{aligned} \overline{\theta(t)\theta(t+\sigma)} &= \overline{\theta(t+\delta t)\theta(t+\delta t+\sigma)} \\ &= \overline{\{\theta(t) + \delta t\dot{\theta}(t)\}\{\theta(t+\sigma) + \delta t\dot{\theta}(t+\sigma)\}} \\ &= \overline{\theta(t)\theta(t+\sigma)} + \delta t\{\overline{\dot{\theta}(t)\theta(t+\sigma)} + \overline{\theta(t)\dot{\theta}(t+\sigma)}\}. \end{aligned}$$

TABLE 1
AUTOCORRELATION COEFFICIENTS

Date	Run No.	Time (E.S.T.)	Height z (m)	\bar{u}^* (m sec ⁻¹)	z/\bar{u} (sec)	Temp. Diff.† 2-32 m (°F)	Time Lag σ (sec)															
							1	1½	2	2½	3	4	5	6	8	10	15	20	30	40		
8.xi.51	1	1241	1.5	2.18	0.69	-0.4	<i>u</i>	0.92	0.81	0.74	0.62	0.52	0.34	0.24	0.18	0.05	-0.10	-0.17				
							<i>v</i>	0.83	0.75	0.67	0.56	0.46	0.38	0.39	0.35							
							<i>w</i>	0.40	0.27	0.17	0.06	0.02	-0.01	0.00								
8.xi.51	2	1251	29	7.85	3.7	-0.9	<i>u</i>	0.94	0.93	0.88	0.87	0.84	0.76	0.68	0.64	0.58	0.57	0.57				
							<i>v</i>	0.93	0.89	0.84	0.77	0.71	0.63	0.54	0.53	0.46						
							<i>w</i>	0.74	0.62	0.52	0.35	0.27	0.09	0.02	0.01							
8.xi.51	3	1300	1.5	4.47	0.34	-0.5	<i>u</i>	0.73	0.63	0.59	0.50	0.41	0.28	0.29	0.20	0.01	0.03	0.06				
							<i>v</i>	0.67	0.34	0.42	0.31	0.26	0.19	0.16	0.17							
							<i>w</i>	0.30	0.17	0.15	0.04	0.01	-0.01	-0.07	0.03							
8.xi.51	4	1308	29	7.70	3.8	-1.0	<i>u</i>	0.92	0.89	0.87	0.85	0.82	0.79	0.75	0.72	0.66	0.60	0.55	0.45			
							<i>v</i>	0.87	0.81	0.76	0.73	0.68	0.60	0.57	0.54	0.47	0.39	0.31				
							<i>w</i>	0.82	0.61	0.49	0.25	0.07	0.01	0.07								
9.i.52	1	1422	29	4.25	6.8	-1.3	<i>u</i>	0.93	0.86	0.80	0.73	0.62	0.49	0.41	0.32	0.26	0.05	0.05				
							<i>v</i>	0.83	0.72	0.64	0.48	0.44	0.35	0.17	0.02	-0.11	-0.08	0.14				
							<i>w</i>	0.95	0.87	0.79	0.73	0.68	0.55	0.37	0.29	0.21	0.16	-0.02				
9.i.52	2	1438	1.5	3.91	0.38	-0.5	<i>u</i>	0.75	0.61	0.51	0.40	0.28	0.11	0.07	0.06	-0.04	-0.03	0.03				
							<i>v</i>	0.83	0.75	0.63	0.55	0.46	0.41	0.33	0.27	0.11	0.09					
							<i>w</i>	0.29	0.15	0.13	0.13	0.14	0.01	0.15	0.14							
9.i.52	3	1549	29	4.60	6.3	-1.0	<i>u</i>	0.81	0.71	0.60	0.47	0.38	0.33	0.28	0.17	0.13	0.23					
							<i>v</i>	0.76	0.70	0.63	0.51	0.42	0.31	0.31	0.39							
							<i>w</i>	0.96	0.93	0.82	0.82	0.80	0.64	0.62	0.48	0.29						
9.i.52	4	1557	1.5	3.82	0.39	-1.6	<i>u</i>	0.81	0.72	0.62	0.51	0.45	0.28	0.12	0.02	-0.04	0.04					
							<i>v</i>	0.81	0.72	0.60	0.46	0.28	0.14	0.09	-0.02	-0.08	0.03					
							<i>w</i>	0.53	0.21	0.18	0.09	0.08	-0.11	0.01								
9.x.51	2	1624	7	4.18	1.7	Not recorded	<i>u</i>	0.81	0.65	0.60	0.57	0.49	0.41	0.37								
							<i>v</i>	0.88	0.83	0.80	0.77	0.77	0.71	0.70								
							<i>w</i>	0.51	0.38	0.19	0.12	0.06	0.07	0.08								

* At height z .

† Negative sign indicates lapse.

Therefore

$$\overline{\dot{\theta}(t)\theta(t+\sigma)} + \overline{\theta(t)\dot{\theta}(t+\sigma)} = 0.$$

By further applications of this principle and also by relating the superior dot (which, of course, denotes d/dt) to the operator $d/d\sigma$ we simplify the expression for $R(\sigma)$ and obtain

$$R = \rho - \frac{2}{\omega^2} \frac{d^2\rho}{d\sigma^2} + \frac{1}{\omega^4} \frac{d^4\rho}{d\sigma^4},$$

so that R can be calculated numerically from a graph showing ρ as a function of σ .

In practice, however, it was found that the terms in $d^2\rho/d\sigma^2$ and $d^4\rho/d\sigma^4$ were small except near $\sigma=0$ where the derivatives of the ρ curve are hard to assess. For $\sigma \geq 1$ sec, it has therefore been taken that $R(\sigma) = \rho(\sigma)$, but there remains the possibility that a significant error may exist in the mean square of u' , that is, in $R(0)$.

From the autocorrelation coefficients and mean square velocity fluctuations were calculated mean square velocity differences

$$D_u(\sigma) = \overline{\{u(t) - u(t+\sigma)\}^2} = 2\overline{u'^2} \left\{ 1 - \frac{R}{u'^2} \right\},$$

and similar functions $D_v(\sigma)$ and $D_w(\sigma)$. The logarithms of D were plotted against those of σ and a typical example of such a graph is shown in Figure 1.

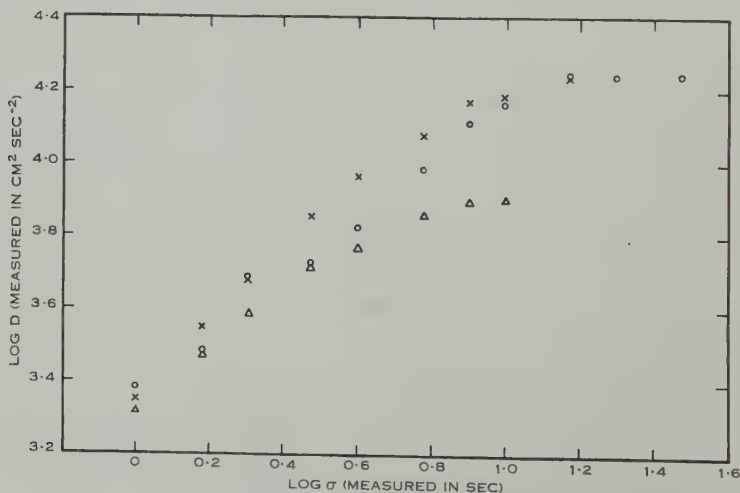


Fig. 1.—Variation of mean square velocity difference with time lag.

$$\bigcirc, (\overline{u-u_\sigma})^2; \times, (\overline{v-v_\sigma})^2; \triangle, (\overline{w-w_\sigma})^2.$$

When $R/\overline{u'^2}$, the correlation coefficient between $u(t)$ and $u(t+\sigma)$, is large, a comparatively small error in $\overline{u'^2}$ will have a disproportionately large effect on the value of D . In some of the graphs of $\log D$ against $\log \sigma$, the values of D for $\sigma=1$ sec and $\sigma=1.5$ sec appear to be rather too small to be consistent with the trend of D for somewhat larger σ . This effect has been attributed to under-

estimation of $\overline{u^2}$ and the points concerned have been ignored in considering the variation of D with σ .

With the reservation expressed in the previous paragraph, every graph showed a fairly straight section rising at slope p , representing the index in the proportionality

$$D \propto \sigma^p,$$

followed by a section of continually decreasing slope and occasionally one of negative slope. In Table 2 are presented the values of p for the straight sections of the curves and the approximate limiting values of σ beyond which the slopes begin to decrease. Values of z/\bar{u} from Table 1 are repeated here so that they may be compared with these limiting σ .

TABLE 2
VARIATION OF MEAN SQUARE VELOCITY DIFFERENCE WITH TIME LAG

Date	8.xi.51			8.xi.51			8.xi.51		
Run No.	1			2			3		
Mean sq. differences in:	<i>u</i>	<i>v</i>	<i>w</i>	<i>u</i>	<i>v</i>	<i>w</i>	<i>u</i>	<i>v</i>	<i>w</i>
p	1.05	0.87	0.51	0.78	1.03	0.83	0.54	0.72	0.19
Limiting σ (sec) ..	4	4	2	10	4	3	6	3	4
z/\bar{u} (sec)	0.69			3.7			0.34		
Date	8.xi.51			9.i.52			9.i.52		
Run No.	4			1			2		
Mean sq. differences in:	<i>u</i>	<i>v</i>	<i>w</i>	<i>u</i>	<i>v</i>	<i>w</i>	<i>u</i>	<i>v</i>	<i>w</i>
p	0.51	0.52	0.90	0.77	0.67	0.72	0.58	0.83	*
Limiting σ (sec) ..	>40	6	4	10	10	10	6	4	<1½
z/\bar{u} (sec)	3.8			6.8			0.38		
Date	9.i.52			9.i.52			9.x.51		
Run No.	3			4			2		
Mean sq. differences in:	<i>u</i>	<i>v</i>	<i>w</i>	<i>u</i>	<i>v</i>	<i>w</i>	<i>u</i>	<i>v</i>	<i>w</i>
p	0.84	0.64	†	0.51	0.97	0.25	0.61	0.77	0.68
Limiting σ (sec) ..	3	4	†	10	4	6	4	3	2½
z/\bar{u} (sec)	6.3			0.39			1.7		

* Indeterminate: not enough points on straight portion of graph.

† Indeterminate: points too badly scattered for reasonable estimate.

It is apparent from Table 2 that the quoted values of p apply up to limiting values of σ which are, in almost every case, several times as large as z/\bar{u} and which occasionally rise to 10 or 20 times as large. This is especially so in the case of u , though unexpectedly large limiting σ for v and w are by no means absent.

In Table 3, means and standard errors of the means of p are given for u , v , and w separately and for all of them taken together. There is some basis for suggesting that the observed difference between the mean p for v and that for w may be significant. If this is so, it is probably to be attributed to the restricting action of the underlying surface on the vertical component and to the relative increase in constraint with increasing scale of motion.

III. COMPARISON WITH THEORY

In considering the dependence of D on parameters of the flow, the theory applicable to the inertial subrange immediately suggests either ε and σ or ε , \bar{u} , and σ (where ε is the rate of dissipation of kinetic energy per unit mass) as possible groups of defining parameters. In laboratory studies of turbulent

TABLE 3
VALUES OF p IN $D \propto \sigma^p$

Component	u	v	w	All
Mean p	0.69	0.78	0.58	0.69
Standard error of mean p .. .	0.06	0.05	0.10	0.04

structure, correlations of this type are invariably taken as equivalent to spatial correlations over a distance $\bar{u}\sigma$ and MacCready (1953) and other authors have extended this assumption of equivalence to the atmosphere. If this is assumed, dimensional analysis indicates that

$$D \propto \varepsilon^{2/3} \bar{u}^{2/3} \sigma^{2/3}.$$

The present writer (1952*a*) previously took the view that D should be considered as a function of ε and σ alone. Robinson (1953) took up an analogous position in considering energy spectra as a function of a time variable (frequency) rather than a space variable (wave number) and he adduced evidence in support of this attitude. If D depends only on ε and σ , then dimensional analysis leads us to

$$D \propto \varepsilon \sigma.$$

However, the observed mean p for all components offers strong support for believing that D is proportional to $\sigma^{2/3}$. The variation of D with ε , \bar{u} , and σ (if only these three are admitted as significant variables) now becomes determinate on dimensional grounds, D being proportional to the two-thirds powers of ε and \bar{u} , as well as of σ .

A further comparison of these observations with the theory can be made by taking advantage of an expression given by Obukhov and Yaglom (1951)

$$6\nu \frac{\partial}{\partial r} \{ \overline{(u-u_r)^2} \} + |S| \{ \overline{(u-u_r)^2} \}^{3/2} = 4\varepsilon r/5,$$

where ν is the kinematic viscosity, u and u_r are the velocity components, directed along r , at two points distant r from one another, and S is the skewness in the distribution of $(u-u_r)$. This equation was given in another form by Batchelor (1947) but Obukhov and Yaglom point out that, in the terms of the theory, S must be an absolute constant within the limits of eddy size to which the theory applies. To make use of this expression we must assume that the mean square velocity differences here considered are equivalent to $\overline{(u-u_r)^2}$ with r given by $\bar{u}\sigma$. Only $D_u(\sigma)$ is of the necessary "longitudinal" form; D_v and D_w are concerned with velocity components at right angles to the distance $\bar{u}\sigma$.

Following Obukhov and Yaglom, we accept a value of -0.4 for S and neglect the term containing v in the inertial subrange. We have then

$$0.4D^{3/2} = 4\varepsilon\bar{u}\sigma/5,$$

and from the constant- p portion of the graphs referred to above it is possible to make an estimate of ε for each occasion.

The present writer (Taylor 1952*b*) has brought forward reasons why ε can be taken as equal to $(\tau/\rho)(\partial\bar{u}/\partial z)$ (where τ is the shearing stress and ρ the density) in steady conditions and on occasions when the heat flux is not too large. From the values of ε obtained as described above, values of τ were thus computed, making use of a wind gradient calculated by logarithmic interpolation between mean wind speeds measured simultaneously at two heights. These were then compared with values of τ determined at the same time from the covariance between vertical and horizontal components. The results of this comparison are displayed in Table 4.

TABLE 4
COMPARISON OF SHEARING STRESSES

Date	Run No.	Shearing Stress (dyn cm ⁻²)	
		From Autocorrelations	From cov(u, w)
8.xi.51	1	0.19	0.92
	2	3.50	1.41
	3	0.49	2.26
	4	*	(-0.86)
9.i.52	1	*	(-0.57)
	2	1.36	1.06
	3	4.31	2.05
	4	0.99	1.81
9.x.51	2	0.24	0.15
Mean..	1.58	1.38

* The existence, on these occasions, of an upward flux of momentum indicates conditions under which the approximation $\varepsilon = (\tau/\rho)(\partial\bar{u}/\partial z)$ cannot be expected to apply. These occasions are therefore omitted from the comparison.

It is worthy of note that there is relatively good agreement, on the average, between the shearing stresses calculated by the two different methods and this agreement offers support for the suggestion that the observed proportionality between D and $(\bar{u}\sigma)^{2/3}$ is, indeed, a continuation of a similar proportionality at smaller eddy sizes where Obukhov and Yaglom's expression can be expected, on solid theoretical grounds, to apply. For the individual runs, it will be noticed that the two shearing stresses often differ quite widely, although the difference is never as great as an order of magnitude.

From the two types of comparison made above, it would appear that properties which might be predicted for the inertial subrange can also apply, on the average, to eddies of down-wind dimension many times the height of observation.

It is, perhaps, not obvious that these eddies are in fact outside the range of local isotropy. Robinson (*loc. cit.*) claims to have shown equipartition of kinetic energy between vertical and horizontal components in the range 3–8 c/s at 1.5 m above the ground but no evidence of equipartition at lower frequencies appears to emerge from his observations. Webb (1955) has carried out some spectrum analyses, and from his results it is possible, on two occasions, to demonstrate the existence of a band of relatively low frequency having approximate equipartition between two components. In neither case, however, does this band appear to be an extension of a region of more exact equipartition at higher frequencies.

Eddies within the range of local isotropy must, moreover, show not only equipartition of kinetic energy but also zero correlation between two velocity components, so that they cannot contribute to the turbulent flux of momentum. The contribution of various eddy sizes to this flux has been studied by Deacon (1955) who concludes, from an analysis of six measurements of momentum flux over 5-min periods, that the greatest contribution to the flux at a height of 28 m is from eddies ranging in period from 5 sec to 3 min. At the average wind speed existing for these measurements, this indicates that the smallest eddy contributing to the momentum flux at this height has a down-wind dimension of at most 40 m. This figure is supported by a result of Panofsky (1953) who gives a stress spectrum analysis for a height of 23 m. At the existing wind speed, the highest frequencies contributing to the shearing stress represent eddies of dimension about 40 m.

Again, a flux measurement at 7 m which Deacon has analysed shows maximum contribution to the flux by eddies of period 16 sec and a small contribution from those of period as short as 1 sec, with a mean wind speed of 3.27 m sec⁻¹. This indicates that eddies of down-wind dimension about 3 m make some contribution to the momentum flux at this height.

In order to obtain lengths from the present results to compare with those quoted, the limiting σ (for autocorrelations in u) as shown in Table 2 were multiplied by the appropriate mean wind speeds. At heights of 1.5 and 29 m, where there are several observations, there is considerable variability in $\bar{u}\sigma_{\text{lim}}$ but the average values are 24 and 110 m respectively. For the one result at 7 m, $\bar{u}\sigma_{\text{lim}}$ is 17 m. Further work on the spectrum of the momentum flux is in progress but it already seems clear that considerations both of this spectrum and of the partition of kinetic energy among components indicate that the observations here dealt with pertain to eddies outside the locally isotropic range.

The continuity relationship between autocorrelation coefficients derived by de Karman and Howarth (1938) can be used as a test of isotropy. The

analogue of this relationship for local isotropy has been stated by Batchelor (1947) and is

$$D_n = D_l + \frac{r}{2} \frac{\partial}{\partial r} D_l,$$

where D_l is a "longitudinal" mean square velocity difference and D_n a "transverse" one between two points separated by r . Here, D_l corresponds to D_w , D_n to D_v or D_w , and $r = \bar{u}\sigma$. By putting $D_l = D_w$, values of D_n were calculated in the form given above and from them were derived values of D_v/D_n and D_w/D_n (which will be equal to unity in the locally isotropic range). There was a considerable amount of scatter but mean values of these ratios at 1.5 m and 29 m are given in Table 5. A mean was not calculated unless there were at least three results to contribute to it.

TABLE 5
TEST OF LOCAL ISOTROPY

σ (sec)	1	$1\frac{1}{2}$	2	3	4	6	8	10	15
D_v/D_n } at 29 m	0.88	0.90	0.89	1.02	0.98	0.92	0.93	0.87	0.98
D_w/D_n }	0.47	0.67	0.85	0.88	0.84	0.94	0.98	1.28	
D_v/D_n } at 1.5 m	0.41	0.42	0.43	0.43	0.47	0.46	0.45	0.42	
D_w/D_n }	0.42	0.35	0.30	0.26	0.24	0.22	0.19		

Only in the case of D_v/D_n at 29 m is the theoretical value at all nearly approached with any consistency though it might be considered (neglecting the somewhat suspect values of D for $\sigma=1$ and $1\frac{1}{2}$ sec) that D_w/D_n at this height is not significantly different from unity. However, it is more likely that the steady growth of D_w/D_n at 29 m from 0.47 to 1.28 is a reflection of the difference between the mean observed p for u and w at this height (0.72 and 0.82 respectively). Evidence in Table 5 for the existence of local isotropy is very doubtful at 29 m and quite absent at 1.5 m, and the previous conclusion, that the observations here dealt with represent eddies outside the locally isotropic range, is supported.

IV. CONCLUSIONS

It is generally accepted that, in the inertial subrange, the rate of viscous dissipation of energy is the only parameter influencing the structure of the flow, the statistical characteristics of which are uniquely determined by it. In the results here analysed, however, this dependence on ε alone does not appear to exist. On the average, predictions of the theory of local isotropy are fulfilled surprisingly well over a range of eddy sizes to which they were never intended to apply but there are considerable discrepancies in individual observations.

This suggests that ε , though it no longer possesses all the power it has in the inertial subrange, remains an important parameter in the range of eddy sizes considered, even though other variables (wind profiles, stability, and the like) are beginning to have an effect. The evidence shows that the rate of viscous dissipation of kinetic energy has a considerable amount of influence in

defining the structure of eddies of a size important in micrometeorological investigations and deserves more attention in this type of work than it has hitherto received.

V. ACKNOWLEDGMENTS

The author's thanks are due to Mr. F. K. Ball for helpful discussions of the effect of galvanometer response time on autocorrelations and to Mr. E. L. Deacon for much general criticism and advice.

VI. REFERENCES

- BATCHELOR, G. K. (1947).—*Proc. Camb. Phil. Soc.* **43**: 533.
DEACON, E. L. (1955).—C.S.I.R.O. Aust. Div. Met. Phys. Tech. Pap. No. 4.
DE KARMAN, T., and HOWARTH, L. (1938).—*Proc. Roy. Soc. A* **164**: 192.
MACCREADY, P. B. (1953).—*J. Met.* **10**: 434.
OBUKHOV, A. M., and YAGLOM, A. M. (1951).—*Appl. Math. Mech., Leningr.*, **15**: 3.
PANOFSKY, H. A. (1953).—Sci. Rep. Penn. St. Coll. Min. Ind. Exp. Sta. No. 2.
ROBINSON, G. D. (1953).—Met. Res. Comm., Lond. MRP No. 808.
TAYLOR, G. I. (1922).—*Proc. Lond. Math. Soc.* **20**: 196.
TAYLOR, R. J. (1952*a*).—Geophys. Res. Pap., M.I.T. No. 19, p. 231.
TAYLOR, R. J. (1952*b*).—*Quart. J. R. Met. Soc.* **78**: 179.
WEBB, E. K. (1955).—C.S.I.R.O. Aust. Div. Met. Phys. Tech. Pap. No. 5.

DISPERSAL OF DUST PARTICLES FROM ELEVATED SOURCES

By G. T. CSANADY*

[Manuscript received June 8, 1955]

Summary

The equations of Sutton for the atmospheric diffusion of gaseous pollutants are extended to the case of particles of appreciable free falling velocity with the boundary condition that dust fall rate equals ground level concentration times free falling velocity. It is shown that the effect of gravitational settling can be allowed for by (a) substituting the height above ground of the downward deflected plume axis for chimney height and (b) by introducing a factor $\alpha(x,z)$ as a multiplier to the "reflection" term in Sutton's equation.

I. INTRODUCTION

Owing to a large extent to the excellent work of Sutton there now exists a rational theory of diffusion of gaseous and finely divided particles from industrial stacks (Sutton 1932, 1947*a*, 1947*b*). Although the theory rests on certain simplifying assumptions, it is sufficiently accurate for practical applications, taking into account, among other factors, the influence of atmospheric stability.

An alternative treatment has been given by Bosanquet and Pearson (1936) and later applied to larger size particles by Bosanquet, Carey, and Halton (1950). This theory, however, fails to account for atmospheric stability and is, as regards the extension to finite size particles at least, less rigorous than Sutton's.

An extension of Sutton's formulae to cover gravitational settling has been attempted by Baron, Gerhard, and Johnstone (1949); an explicit solution has however, not been obtained.

From an engineering point of view rates of dust deposition are probably at least as important as maximum ground level concentrations of gaseous pollutants, the latter constituting a "public nuisance" relatively rarely while communal protests on grit fall have become increasingly numerous in the vicinity of major industries. The present article is an attempt to apply Sutton's theories to the problem of dust deposition.

II. SUTTON'S EQUATIONS

Using a theorem due to Taylor (1922), Sutton (1932) has obtained an expression for the standard deviation of diffusing particles from their mean position

$$\sigma^2 = \frac{1}{2} C^2 x^{2-n},$$

* New South Wales University of Technology, Wollongong College.

where σ is the root-mean-square distance of particles, e.g. in a cloud of smoke from the centre of the cloud, C is the (constant) virtual eddy diffusion coefficient, and $x=ut$ the distance travelled down-wind by the cloud in a wind of mean velocity u in time interval t . The index " n " is a parameter affected by atmospheric stability, having the value 0.25 in a neutral atmosphere, 0-0.25 for high lapse rates, and 0.25-1.0 for inversions.

With the aid of this result Sutton (1947*a*, 1947*b*) has derived a good approximate expression for eddy diffusion of gases emitted by a chimney:

$$c = \frac{W}{\pi C_y C_z u x^{2-n}} \exp \left(-\frac{y^2}{C_y^2 x^{2-n}} \right) \left[\exp \left\{ -\frac{(z-h)^2}{C_z^2 x^{2-n}} \right\} + \exp \left\{ -\frac{(z+h)^2}{C_z^2 x^{2-n}} \right\} \right],$$

where

c = concentration of pollutant at some point at a distance x down-wind from the source, a perpendicular horizontal distance y from the plume axis, and an elevation z above ground level;

W = strength of source, emitted matter per unit time;

C_y, C_z = virtual diffusion coefficients in the y and z directions respectively;

h = chimney height;

u = mean wind velocity;

n = parameter of stability.

On close examination the expression in the square brackets is seen to consist of two terms with a definite physical significance: the first one gives the diffusion from a plume at a mean distance h above ground as if the ground were not present. The second one is the exact mirror image of the first and has the effect of exactly replacing that part of the plume "cut off" by the ground; it may be regarded as the "reflected" plume. By this device the condition of continuity, namely, that all emitted particles remain airborne, is satisfied.

III. EFFECT OF GRAVITY ON PLUME

If now the emitted particles possess a groundward velocity not negligible compared to the wind velocity, it may be assumed with confidence that their mean motion will be the geometrical sum of windward travel and free fall, i.e. that the axis of the plume will have an inclination to the horizontal equal to f/u where f is the velocity of free fall for the particles concerned. Or more accurately, if z^* denotes the vertical distance of the plume axis above ground level at a distance x down-wind from the chimney, then

$$z^* = h(1 - x/x_0),$$

where x_0 is the point at which the inclined plume axis reaches the ground. Also (Fig. 1):

$$x_0 = (u/f)h,$$

therefore

$$z^* = h\{1 - (f/u) \cdot (x/h)\}.$$

In Sutton's equation the mean height of the particles above ground was $h = \text{const}$; replacing this by the variable z^* ,

$$c = \frac{W}{\pi C_y C_z u x^{2-n}} \exp \left(-\frac{y^2}{C_y^2 x^{2-n}} \right) \times \left[\exp \left(-\frac{\{z + (f/u)x - h\}^2}{C_z^2 x^{2-n}} \right) + \exp \left(-\frac{\{z - (f/u)x + h\}^2}{C_z^2 x^{2-n}} \right) \right].$$

This equation now gives a deflected plume path but the presence of the mirror image term still guarantees that all particles are "reflected" by the ground. Physically this is incorrect because in the stagnant air layer immediately adjacent to the ground the motion of the particles consists solely

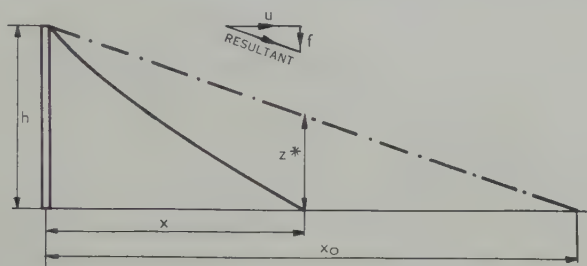


Fig. 1.—Groundward deflection of plume due to free fall of particles; mean velocity is given by geometrical sum of wind velocity and fall velocity. "Streamlines" in random diffusion are given by the condition that distance from the mean position is a constant multiple of the local standard deviation.

of free fall; and, as the concentration in this layer must be substantially the same as in the bordering turbulent mass, a transport of particles given by $c(z=0) \cdot f$ settles through the stagnant layer to the ground. Thus the continuity condition, one of the boundary conditions for Sutton's solution, is replaced by the prescription that the rate of dust deposition at any point on the ground should be $c_0 f$ and that the airborne particles should be diminished by a corresponding amount.

IV. ALLOWANCE FOR DEPOSITION

If the mirror image term in the last equation were neglected altogether, that is, if it were assumed that all particles reaching the ground by diffusion would be retained there, none reflected, the effective settling velocity of particles would be given by the sum of the mean vertical velocity of particles, i.e. f , and the instantaneous rate of increase of their mean distance from the plume axis. Let this latter term be called Δw ; then from the foregoing it is clear that, the vertical distribution of particles remaining "normal" with increase of x , Δw will be proportional to $d\sigma_z/dt$, moreover, the factor of proportionality will be z^*/σ_z , i.e. the vertical distance of the plume axis above ground expressed in terms of standard deviations. To see this, consider that a constant fraction of

particles is enclosed by the surface $\rho = \text{const.}$ σ (where $\rho = \text{distance from plume axis}$) and that therefore mean flow is along this surface. Consequently,

$$\Delta w = \frac{z^*}{\sigma_z} \cdot \frac{d\sigma_z}{dt}.$$

σ_z may be found with Sutton,

$$\sigma_z = \frac{1}{2} \sqrt{2C_z x^{1-\frac{1}{2}n}},$$

from which it follows that

$$\Delta w = (1 - \frac{1}{2}n)(ku/x - f).$$

The significance of this result is that if the mirror image term is neglected the condition of continuity requires a dust fall rate of $D = c_0(f + \Delta w)$ to be postulated instead of the physically correct $c_0 f$.

On the other hand, it is not necessary to neglect the whole of the mirror image term: a fraction may be retained so that the dust fall rate is exactly as required.

It should be noted now that in computing ground level concentrations (c_0) the mirror image term reduces to a form identical to the simple diffusion term, i.e. that the mirror image term doubles the ground level concentration obtained by the simple term alone. Also, the upward directed "settling velocity" contributed by the mirror image term is numerically equal but opposite in sense

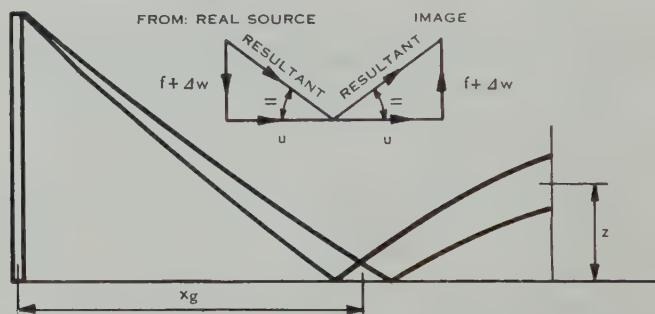


Fig. 2.—"Reflection" of plume by ground. Strength of "reflected" beam has to be adjusted to satisfy boundary condition that settling rate equals ground level concentration multiplied by free fall velocity.

to the velocity of settling from the simple plume (Fig. 2). If now the reflection term is multiplied by a factor $\alpha_0(x)$ the net rate of settling may be postulated to be the physically correct one

$$c_{s_0}(1 + \alpha_0)f = c_{s_0}(f + \Delta w) - \alpha_0 c_{s_0}(f + \Delta w),$$

where c_{s_0} is the ground level concentration given by the simple diffusion term alone and $\alpha_0 c_{s_0}$ the contribution from the retained fraction of the reflection term.

From the above

$$\alpha_0(x) = \Delta w / (2f + \Delta w),$$

which, with the previously obtained value of Δw , reduces to

$$\alpha_0(x) = 1 - \frac{2}{(1-n/2)(hu/xf-1)+2}.$$

The value of α_0 thus computed refers to the degree of reflection on the ground which is seen to vary with distance from the source. Above ground level the strength of the reflected beam will be governed by the value of α_0 at the point of reflection (Fig. 2),

$$\alpha(z, x) = \alpha_0(x_g).$$

The computation of $\alpha(z, x)$ unfortunately presents a certain amount of difficulty owing to the fractional powers of x that have to be dealt with. As previously remarked, the equation for "streamlines" becomes:

$$z = z^* - k\sigma, \quad k = \text{const.}$$

or, using Sutton's expression for σ and the equation of z^* derived above,

$$z = h - (f/u)x - \frac{1}{2}k\sqrt{2C_z}x^{1-\frac{1}{2}n}$$

for $z \rightarrow 0, x \rightarrow x_g$ (Fig. 2).

Therefore

$$\frac{1}{2}k\sqrt{2C_z} = \{h - (f/u)x_g\}/x_g^{1-\frac{1}{2}n},$$

and with this

$$z = h - (f/u)x - \{h - (f/u)x_g\}(x/x_g)^{1-\frac{1}{2}n}.$$

The significance of the last equation is that, given an arbitrary point (x_g, z) , the point of "reflection"—which by a previously derived relationship determines the strength of the "reflected" beam—may be found by solving this equation for x_g . A general explicit solution is not readily obtained so that $\alpha(x, z) \equiv \alpha_0(x_g)$ cannot be given in an explicit form. This fact should not be of very great practical disadvantage as concentrations above ground level are mostly of academic interest only. If, in experimental checks on the theory, measurements above ground level are conducted, predictions of the theory have to be found by the somewhat cumbersome method of first finding x_g numerically. In such cases, however, the complication can be tolerated.

Modifying now Sutton's expressions by inclusion of the multiplier α the solution for concentration of particles at any point (x, y, z) becomes

$$c = \frac{W}{\pi C_y C_z u x^{2-n}} \exp \left(-\frac{y^2}{C_y^2 x^{2-n}} \right) \times \left[\exp \left(-\frac{\{z + (f/u)x - h\}^2}{C_z^2 x^{2-n}} \right) + \alpha \exp \left(-\frac{\{z - (f/u)x + h\}^2}{C_z^2 x^{2-n}} \right) \right],$$

where

$$\alpha = 1 - \frac{2}{(1-\frac{1}{2}n)(uh/x_g f - 1) + 2},$$

$x_g = x_g(z, x)$ to be found from

$$\{h - (f/u)x_g\}(x/x_g)^{1-\frac{1}{2}n} + (f/u)x + z - h = 0.$$

Ground level concentrations can be expressed explicitly as $x_g=x$; dust fall rates are then c_0f , or

$$D = \frac{Wf}{\pi C_y C_z u x^{2-n}} \exp \left(-\frac{y^2}{C_y^2 x^{2-n}} - \frac{\{(f/u)x-h\}^2}{C_z^2 x^{2-n}} \right) \left[2 - \frac{2}{(1-\frac{1}{2}n)(uh/xf-1)+2} \right].$$

Regarding units it should be noted that, as pointed out by Sutton (1947b), C_y and C_z have the dimension of $L^{1/n}$. The dimension of D will then be $ML^{-2}T^{-1}$.

The manner in which an expression for α at $z \neq 0$ was obtained should guarantee that continuity is preserved, i.e. that the rate of disappearance of particles from suspension equals the rate of their deposition. Unfortunately, the mathematical expression of this condition yields a rather intractable equation

$$\frac{\partial}{\partial x} \int_0^\infty \int_{-\infty}^{+\infty} c u dy dz = \int_{-\infty}^{+\infty} c_0 f dy,$$

which after integration with respect to y reduces to

$$\frac{\partial}{\partial x} \int_0^\infty \frac{u}{x^{3-3n/2}} [\varphi_1(x, z) + \varphi_2(x, z)] dz = [\varphi_1(x, 0) + \varphi_2(x, 0)] f,$$

with

$$\begin{aligned} \varphi_1 &= \exp \left(-\frac{\{z + (f/u)x - h\}^2}{C_z^2 x^{2-n}} \right), \\ \varphi_2 &= \left[1 - \frac{2}{(1-\frac{1}{2}n)(uh/x_g f - 1) + 2} \right] \exp \left[-\frac{\{z - (f/u)x + h\}^2}{C_z^2 x^{2-n}} \right], \\ x_g &= x_g(x, z) \end{aligned}$$

as before. Verification of this equation would appear impracticable by analytical methods.

V. CONCLUSION

Application of Sutton's theories to particles of appreciable size is thus seen to be possible without introducing restricting assumptions, explicit expressions being, however, only obtained for ground level concentrations and dust fall rates, not for concentrations above ground level. The reliability of the formulae arrived at in this paper should not differ very greatly from that of the Sutton equation; they are only intended, however, to be practically applicable approximations.

VI. REFERENCES

- BARON, T., GERHARD, E. R., and JOHNSTONE, H. F. (1949).—Dissemination of aerosol particles from stacks. *Ind. Eng. Chem.* **41**: 2403–8.
- BOSANQUET, C. H., and PEARSON, J. L. (1936).—*Trans. Faraday Soc.* **32**: 1249–63.
- BOSANQUET, C. H., CAREY, W. F., and HALTON, E. M. (1950).—Dust deposition from chimney stacks. *Proc. Inst. Mech. Eng.* **162**: 355–65.
- SUTTON, O. G. (1932).—The theory of eddy diffusion in the atmosphere. *Proc. Roy. Soc. A* **135**: 143–65.
- SUTTON, O. G. (1947a).—*Quart. J. R. Met. Soc.* **73**: 257.
- SUTTON, O. G. (1947b).—The theoretical distribution of airborne pollution from factory chimneys. *Quart. J. R. Met. Soc.* **73**: 426–36.
- TAYLOR, G. I. (1922).—On the theory of diffusion by continuous movements. *Proc. Lond. Math. Soc.* **20**: 196–212.

SHORT COMMUNICATIONS

THE DIFFUSION OF ELECTRONS IN A GAS AT LOW TEMPERATURES*

By BARBARA I. H. HALL†

Introduction

In the absence of an electric field, the state of motion of electrons moving freely among molecules of a gas is one in which the mean energy of agitation Q of an electron is the same as that, Q_0 , of a molecule of the gas. It is known, however, that in the presence of an applied uniform electric field Z the steady state of motion is one in which Q exceeds Q_0 , and, for a constant gas temperature, Q is a function of Z/n where n is the number of gas molecules in unit volume. Usually the measurements are made at room temperature, generally taken as 288 °K, and it is convenient to replace Z/n by Z/p where p is the pressure of the gas in millimetres of mercury. It is of interest, however (Crompton, Huxley, and Sutton 1953), to investigate how the dependence of Q upon Z/n is influenced by the temperature of the gas itself, that is to say, by Q_0 . If the dependence of Q upon both Z/n and Q_0 be expressed by some formula $Q=f(Z/n, Q_0)$, then all that is known with certainty about the function f is that when $Z/n=0$, $Q=Q_0$. In view of the absence of any theoretical guidance concerning the dependence of Q upon both Z/n and Q_0 , it was decided to investigate the matter experimentally in two simple diatomic gases, nitrogen and hydrogen.

Apparatus

The diffusion chamber, of length h cm, was the same as that used by Crompton and Sutton (1952), as were the electrometers and current measuring devices. The chamber was mounted near the bottom of a glass envelope which stood in a cylinder containing either dry ice or liquid air.

The envelope was made considerably longer than that used previously as it was necessary to keep the electrometer terminals at the top as nearly as possible at room temperature in order to avoid deterioration of the electrostatic insulation by the condensation of water vapour. In addition, heater wires were wrapped around the envelope near the top. The leads from the diffusion chamber to the top of the envelope were made of 50 gauge copper wire to reduce the thermal conduction.

The temperature in the diffusion chamber was measured with a copper-constantan thermocouple made of 45 gauge wire, the cold junction being soldered to the earthed ring surrounding the collecting electrode. The hot junction

* Manuscript received June 8, 1955.

† Department of Physics, University of Adelaide.

was at room temperature and the thermo-e.m.f. was measured with a galvanometer and scale. The thermocouple was then calibrated over the useful range of temperatures by using a pentane in nitrogen thermometer as standard. It was assumed that the temperature of the gas in which the electrons were diffusing was the same as that of the lower electrode.

Both the nitrogen and the hydrogen were introduced into the apparatus according to the procedure described by Crompton and Sutton (1952).

Results

The ratio Q/Q_0 is denoted by k_T , but the quantity measured in the diffusion experiments is k_1 where $k_1 = Ak_T$, the factor A being a constant determined by the manner in which the speeds of agitation of the electrons are distributed about their mean value. As described in papers by Huxley and Zaazou (1949) and Crompton and Sutton (1952), the measurement of the ratio of the current falling on the inner circular electrode to that falling on the whole electrode enables k_1 to be determined. Good agreement was obtained between values of k_1 measured at different pressures, but at constant temperature, for the same value of the ratio Z/p .

If k_1 is measured at a temperature T , then

$$Q = \frac{C_M}{C_D} \left\} k_1 T/288 \text{ ergs,} \right.$$

where $C_M = 6.0 \times 10^{-14}$ and $C_D = 5.3 \times 10^{-14}$ are constants appropriate to the distributions of Maxwell and Druyvesteyn respectively.

In previous work in this field, the electronic energy Q is represented as a function of the ratio Z/p where p is the pressure of the gas at 15 °C. It is, however, not the physical pressure p that is significant but the concentration n of the gas molecules per cubic centimetre, of which p is a measure provided the temperature remains constant. Since at 15 °C the concentration of the molecules of a gas exerting a pressure of 1 mm of mercury is $n = 3.35 \times 10^{16} \text{ cm}^{-3}$, the parameter Z/p is in effect the same as $(Z/n) \times 3.35 \times 10^{16}$. When experiments are performed at other temperatures, the pressures p corresponding to the same n differ, and thus in Figures 1 and 2, in order to simplify comparison with other work in this field, the electronic energy Q is represented as a function of $(Z/n) \times 3.35 \times 10^{16}$, which reduces to Z/p at 15 °C. When this is done it is found that, over a large range of values of Z/n , Q is a function of Z/n only and independent of the thermal energies Q_0 of the gas molecules.

In Figures 1 and 2 the curves drawn show the results obtained by Crompton and Sutton (1952), while the points plotted represent the mean energy of agitation of the electrons at -190 °C in nitrogen and -192 °C in hydrogen. Measurements of electron energies made at dry ice temperatures gave essentially the same results.

Conclusions

Since as $Z \rightarrow 0$ the mean energy of agitation Q of an electron approaches Q_0 , it follows that the curve of Q as a function of Z/n for a particular temperature T must cut the axis $Z/n = 0$ at a point corresponding to the thermal energy of

agitation of the gas molecules at that temperature. When a field Z is applied the electrons acquire energy from the field and a value of Z/n is very soon reached at which Q depends only upon Z/n and is independent of the temperature T .

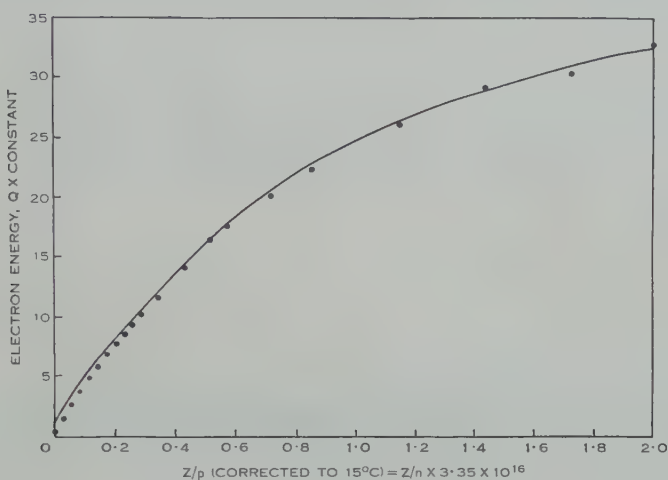


Fig. 1.—Electron energies in nitrogen.

Thus for a constant electric field and fixed molecular concentration the electron energy is independent of the energy of the gas molecules; consequently the motion of the molecules can have very little effect on the diffusion of electrons

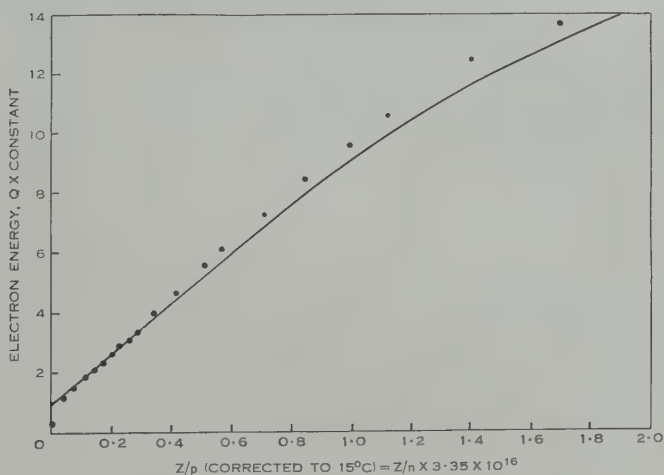


Fig. 2.—Electron energies in hydrogen.

through the gas, and to a first approximation they may be considered to be at rest.

In hydrogen the two sets of results begin to diverge slightly for larger values of Z/n , but it is impossible to say whether or not this is of significance.

The author wishes to thank Professor Huxley for his advice and encouragement and Dr. R. W. Crompton for his guidance and assistance throughout. She would also like to thank the Commonwealth Scientific and Industrial Research Organization for a studentship during the tenure of which this work was carried out.

References

- CROMPTON, R. W., HUXLEY, L. G. H., and SUTTON, D. J. (1953).—*Proc. Roy. Soc. A* **218** : 507.
CROMPTON, R. W., and SUTTON, D. J. (1952).—*Proc. Roy. Soc. A* **215** : 467.
HUXLEY, L. G. H., and ZAAZOU, A. A. (1949).—*Proc. Roy. Soc. A* **196** : 402.

LUNAR TIDE IN SPORADIC *E* AT BRISBANE*

By J. A. THOMAS† and A. C. SVENSON†

Accurate measurements have been made at Brisbane of the virtual height of reflection of radio echoes from the *E* region of the ionosphere, using a pulse transmitter operating at 2.28 Mc/s. The recordings were made by using a cathode-ray tube displaying echoes received from virtual heights between 80 and 150 km; black-out modulation of the cathode-ray tube trace was employed so that echoes appeared as gaps in the trace. To avoid the broadening of the gap with increase of echo signal strength, an automatic gain control was used, which kept the peak output within the time interval corresponding to 80–150 km at a fixed level. The gap width then remained fairly constant (about 16 km), and the recorded height showed no dependence on (input) signal strength. A stable triggered oscillator was used to produce height marks at 10 km intervals every 6 min on the cathode-ray display. The oscillator was started at full amplitude by a triggering wave synchronized with the transmitter modulation pulse; there was a constant delay of 0.1 km. A typical record is shown in Plate 1.

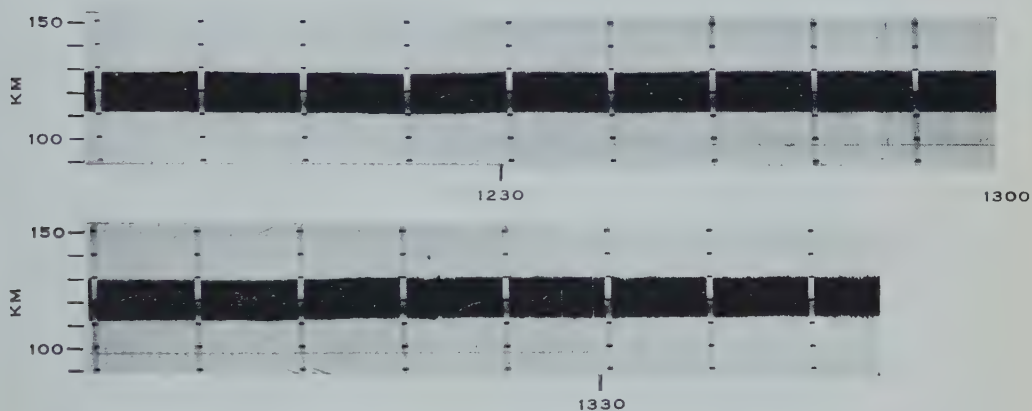
With these precautions it has been possible to take measurements of the virtual height of *E* reflections with an accuracy of ± 0.2 km. It has been found, however, that the levels of both the normal *E* and sporadic *E* reflections vary in a random manner with a quasi-period of about 15 min and a most probable amplitude of about 1 km. This result is to be expected for sporadic *E* reflections and is now well established for the normal *E* region. For this reason there is generally little point in making height measurements to any accuracy greater than ± 0.5 km.

These accurate height measurements lend themselves to lunar tidal analysis. However, at Brisbane the day-time *E* region echo at 2.28 Mc/s may (and generally does) consist of a mixture of normal *E* and sporadic *E* reflections. Sometimes

* Manuscript received August 24, 1955.

† Physics Department, University of Queensland, Brisbane.

LUNAR TIDE IN SPORADIC E AT BRISBANE



Example of $P't$ for steady conditions in E region, using A.G.C. receiver, September 14, 1953.

they may be distinguished by examining simultaneous swept-frequency ($p'f$) records, but only 15 per cent. of all records between 1000 and 1400 hr at this frequency can be positively identified as normal E region. The lunar tidal analysis of $h'E$ performed by Martyn (1948) is thus actually an analysis of a mixture of $h'E$ and $h'E_s$, with probably an emphasis on $h'E_s$. At night-time

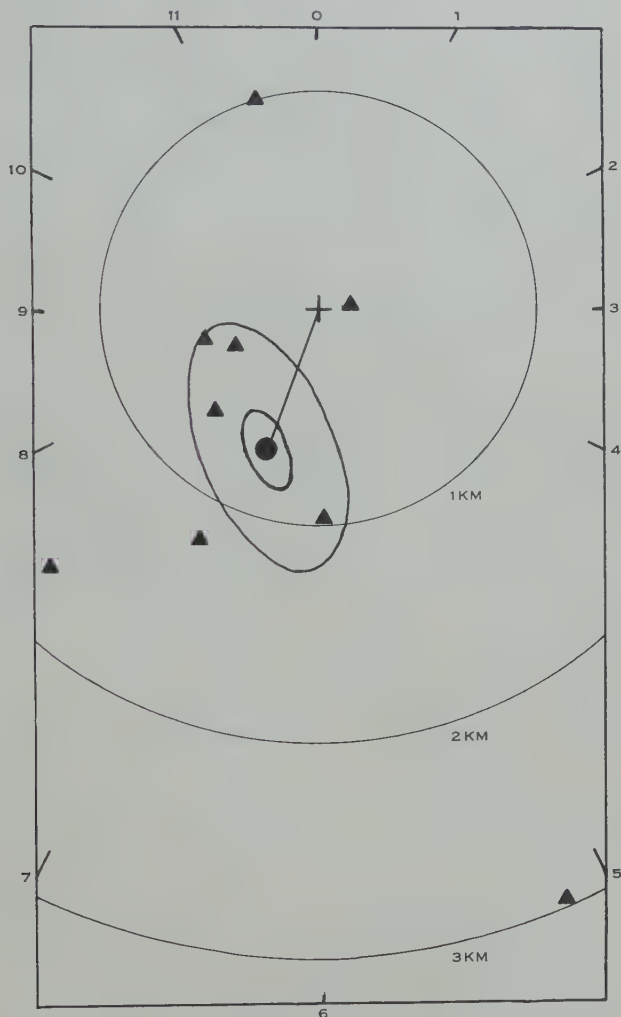


Fig. 1.—Harmonic dial for lunar semi-diurnal tide in $h'E_s$ at Brisbane. \blacktriangle Individual months, \bullet mean.

all echoes are from sporadic E . There may be more than one trace appearing simultaneously on the records, but when this is the case the upper trace generally shows a slope and curvature indicative of decreasing or increasing slant range due to horizontal movements of irregularities. Such traces have been ignored for the purposes of tidal analysis, as also have the saucer-shaped traces ascribed to moving "clouds" of ionization.

A lunar tidal analysis has been made of night-time values of $h'E_s$ for the year December 1952 to November 1953. Measurements were taken at 6 min intervals throughout each night. The results are shown in Table 1, and in diagrammatic form are plotted on the 12 hr harmonic dial of Figure 1. The mean value for the semi-diurnal lunar tide in $h'E_s$ is $P_2=0.69$ km, $t_2=6.7$ hr. Two probable error ellipses are shown. The outermost indicates the probable error of one monthly observation, and the inner ellipse the probable error of the

TABLE 1
VALUES OF AMPLITUDE (P_2) AND PHASE (t_2) OF LUNAR SEMI-DIURNAL TIDES IN NIGHT-TIME VALUES OF $h'E_s$ AT BRISBANE

	Dec.	Jan.	Mar.- Apr.	June	July	Aug.	Sept.	Oct.	Nov.	Mean
P_2 (km)	0.14	0.67	0.96	1.69	0.42	1.00	2.96	0.55	1.20	0.69
t_2 (hr) ..	2.8	7.5	6.0	7.5	8.3	11.4	5.6	8.5	6.9	6.7

computed mean. These indicate that the result is valid, and that changes in the 12-hourly component affect the phase rather than the amplitude of the tide. The mean E_s level throughout the year was 109 km.

Matsushita (1953) has made a similar analysis for Brisbane (and other stations) using the published values of $h'E_s$ determined from $P'f$ records. A comparison of the values obtained is given in Table 2, together with the value obtained by Martyn (1953).

TABLE 2
COMPARISON OF VALUES OF LUNAR TIDES IN E REGION AT BRISBANE

	Type of Records Analysed	P_2 (km)	t_2 (hr)	Period Analysed
Matsushita ..	$h'E_s$ from $P'f$ records (nearest 10 km)	0.70	5.7	May 1951-Apr. 1952
Thomas & Svenson	Night-time $h'E_s$ from $P't$ records (nearest km)	0.69	6.7	Dec. 1952-Nov. 1953
Martyn	$h'E$ from $P'f$ records (nearest 10 km) (Mixture of $h'E$ & $h'E_s$)	0.5	4.5	June 1943-Dec. 1947

The authors are indebted to Professor H. C. Webster and Mr. R. W. E. McNicol for advice and encouragement throughout the course of this work.

The work described forms part of the programme of the Radio Research Board of C.S.I.R.O., and is published by permission of the Board.

References

- MARTYN, D. F. (1948).—*Proc. Roy. Soc. A* **194**: 429.
MATSUSHITA, S. (1953).—*Rep. Ionosphere Res. Japan* **7**: 45.

ELECTRON BACKGROUND IN IMPORTED G5 EMULSIONS*

By V. D. HOPPER† and JEAN E. LABY†

Ilford G5 nuclear emulsions have been received by this department at regular fortnightly intervals during the last six months. They are sent by air mail and reach this laboratory approximately two weeks after the date of manufacture in England and are developed soon after arrival. Heavy electron backgrounds were observed in four batches of plates out of 12 successive fortnightly shipments. Possible causes of these variations have been investigated.

To compare the background in the plates two methods were used which gave closely similar results. In one method the number of sharply defined black grains in a $(50\ \mu)^2$ field using a $95\times$ objective was counted and an average obtained for a number of fields of view. Results are shown in Figure 1 (a). In the second method the number of electron tracks ending in the emulsion in the $(100\ \mu)^2$ field using a $45\times$ objective was counted and averaged for a number of fields of view (Fig. 1 (b)). Neither method is very accurate but the results were quite consistent up to exposures of about 150 milliröntgens. The second method could not be used for exposures above 150 mr as the electron tracks could not be resolved. A comparison was made of these results with plates exposed to a 4.6 c iridium source specified as giving 3 mr/hr at 1 m so as to obtain a measure in milliröntgens of the various backgrounds. The radiation 1 m from the source was checked using a radiation monitor and found to agree with the specified 3 mr/hr. Two Ilford G5 600 μ plates (manufactured May 3) of low background were placed on the surface of the container (i.e. about $\frac{1}{3}$ m from the source) and given exposures of 64 and 113 hr respectively, and two similar plates were placed 2 m from the source and given the same exposure times. The estimated dosages received by various batches are given in Table 1 and plotted in Figures 1 (a) and 1 (b) and photographs showing relative electron backgrounds of the plates manufactured on January 21 and February 7 are shown in Plate 1.

By considering only the cosmic ray background and general radiation from the Earth a rough estimate can be made of the radiation that one would expect in a plate imported to Australia by aircraft and developed one week after arrival, i.e. about a fortnight after manufacture. A typical time of flight in the height region 10,000–20,000 ft would be about 38 hr and the plates would be airborne for 46 hr. Assuming 1 mrep/day at 20,000 ft, the exposure due to cosmic radiation during flight would be of the order of 1.5 mrep. Radiation received at ground level mainly due to γ -rays is of the order of 0.5 mrep/day. Plates are normally stored under 2–3 in. of lead at this laboratory so the rate would be much less than this during storage. Thus the total radiation received

* Manuscript received August 29, 1955.

† Physics Department, University of Melbourne.

by the plates at development would be expected to be about 10 mrep and it will be seen from Table 1 that some of the batches of plates show slow electron backgrounds corresponding to this.

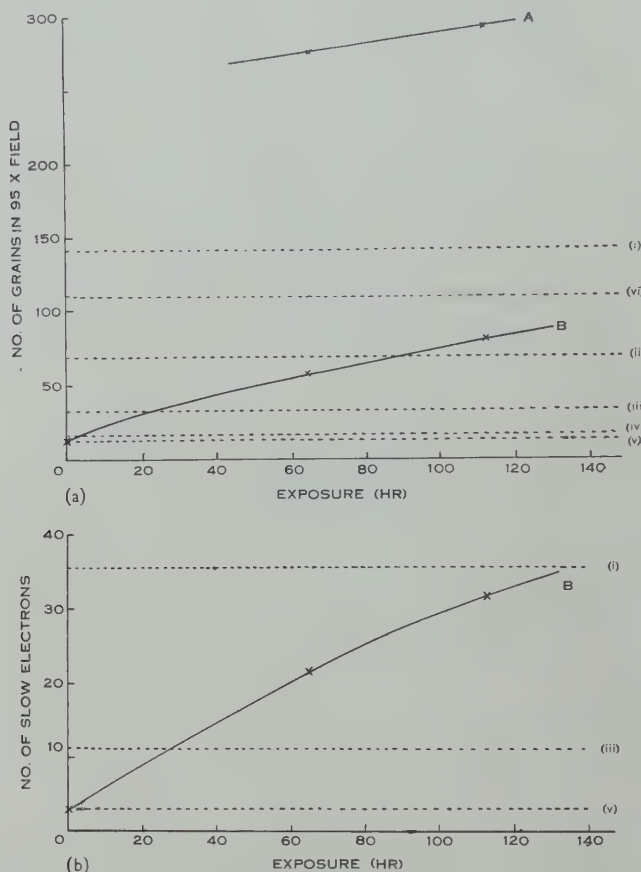


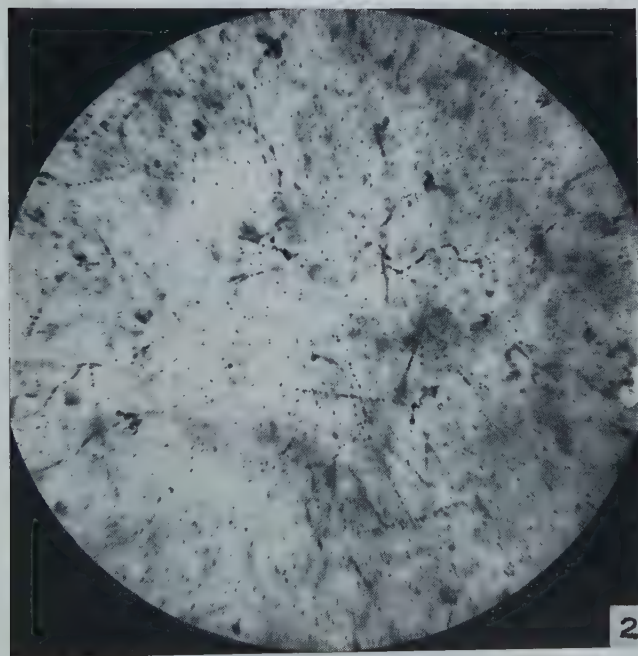
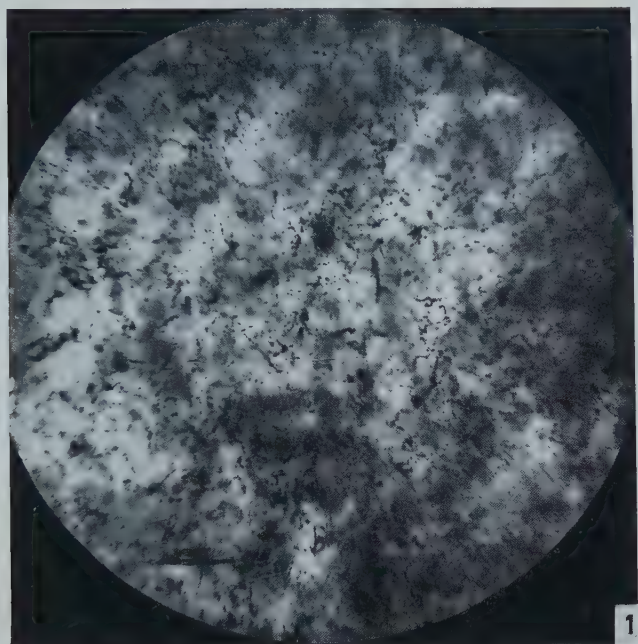
Fig. 1.—Electron background in plates. (a) Number of focused grains in $95\times$ objective field as a function of time of exposure of plates. A, Ilford G5 emulsion exposed $\frac{1}{8}$ m from ^{192}Ir source; B, Ilford G5 emulsion exposed 2 m from ^{192}Ir source. For comparison the level of the background in Ilford G5 plates manufactured: (i) Apr. 1, 1955; (ii) Jan. 21, 1955; (iii) Apr. 20, 1955; (iv) Feb. 7, 1955; (v) May 3, 1955; (vi) May 31, 1955, are shown.

(b) Number of slow electrons ending in $45\times$ field as a function of time of exposure.

Possible causes of the extra radiation background on some of the plates are :

- (a) Irradiation at time of manufacture or during storage in this laboratory.
- (b) Exposure to radiation from instrument panels and radioactive isotopes carried on the aircraft.
- (c) X-ray examination by customs officials.
- (d) Radiation from products of atomic explosion tests.

ELECTRON BACKGROUND IN IMPORTED G5 EMULSIONS



Photographs showing electron background in plates manufactured
(1) Jan. 21, 1955; (2) Feb. 7, 1955.

(a) Irradiation at Manufacture or Storage

It was stated by Ilford that there was no possibility of the emulsion being exposed before dispatch and that occasional complaints that had been received from America were attributed to customs activities or to emulsion travelling in close proximity to radioactive material. Regarding storage at this laboratory, the plates with high background had been stored alongside other plates and these showed no extra background.

TABLE I
DOSAGE MEASUREMENTS ON PLATES OBTAINED BY COMPARING THEIR
BACKGROUND WITH THAT PRODUCED BY γ -RADIATION FROM ^{192}Ir

Date of Manufacture 1955	Departure from London	Estimated Dose on Arrival (mr)
Jan. 21	Jan. 23	70
Feb. 7		15
Apr. 1	Apr. 6 (BOAC)	140
Apr. 20	Apr. 22 (KLM)	35
May 3		10
May 16		10
May 31, June 2	June 4 (KLM)	120
June 13		15
June 28		10

*(b) Contamination in Aircraft**Instrument panels*

A source of contamination in aircraft is the radioactive paint on instrument panels. Measurements were made by Morrison* in 1951 on the γ -radiation from instruments in a "Constellation" aircraft. From his figures it appears that radiation in the luggage compartments presumably well away from the instrument panels would be less than 2 mr per day. As the plates would be stored in the aircraft for the complete flight, the plates might conceivably receive a radiation dosage of 10 mr during the flight due to this cause alone, and it thus seems unlikely that radiation from instrument panels could be as high as the background observed in three of the batches received.

Imported isotopes

The background observed on the two most heavily exposed plates could have been produced if the plates had been stored 4–5 ft† from a 4.6 c ^{192}Ir

* MORRISON, A. (1951).—*J. Aviat. Med.* **22**: 350.

† Boxes of emulsions are labelled specifying that packets should be stored at least 10 ft from any radioactive material and parcels of radioactive isotopes are also labelled advising that photographic material be stored at least 10 ft from the container. Assurances have been received that these instructions are carried out. Ilford G5 emulsions having about 25 mr are usable for cosmic ray experiments but it is of course preferable to have a much lower background. To obtain a dosage less than 25 mr plates imported to Australia should, in fact, be stored at least 15 ft from an iridium source of 4.6 c.

source for four days (see Fig. 1). The Commonwealth X-ray and Radium Laboratory keeps a complete record of all isotopes entering Australia and γ -ray sources of highest activity arriving during the period in which the 12 batches

TABLE 2
RADIOACTIVITY OBSERVED AT VAL-JOYEUX BY Mlle. OLGA TANAEVSKY
AND M. ETIENNE VASSY

Type of Precipitation	Date	Initial Activity (counts $m^{-1} l^{-1}$)
Rain	Jan. 11	61
Snow	Jan. 13	25
Snow	Feb. 15	50
Rain	Feb. 21	8
Rain	Night Feb. 23 to 24	14
Snow	Mar. 7	13
Rain	Night Mar. 20 to 21	88
Rain	Mar. 22	35
Rain	Mar. 28	90
Rain	Night Apr. 6 to 7	1596
Rain	Apr. 10	189
Rain	May 5	112
Rain	May 6	47
Rain	Night May 13 to 14	47
Rain	May 15	87
Rain	Night May 16 to 17	15
Rain	May 18	53
Rain	May 19	20
Rain	May 21*	24
Rain	May 26	69
Rain	June 3	192
Rain	June 6	154
Rain	June 9, 10 hr	87
Rain	18 hr	28
Rain	June 12	37
Rain	June 19	26
Rain	June 30	45
Rain	July 4	38
Rain	July 11	58
Rain	July 18	47
Rain	July 20	26
Rain	July 27	29
Rain	Aug. 10	37
Rain	Aug. 12	17
Rain	Aug. 14	38

* Data for May 21 and later were received in a further personal communication from Professor Vassy and have been added *in proof*.

of plates were received were in fact ^{192}Ir sources of 4.5–5.1 c. The nuclear emulsion plates that had excessive exposures were manufactured on January 21, April 1, April 20, and May 31 and the dates of departure of these batches are given in Table 1. From the records it was found that not one of these batches

travelled to Australia on the same plane as an imported γ -ray isotope. There is, however, the possibility that the plates travelled in planes carrying isotopes dispatched to other countries en route but, if so, the plates would need to be stored even closer to the isotopes than 4.5 ft.

(c) *X-ray Examination at Customs Inspection*

All packets of plates are labelled as containing photographic material and the most heavily exposed was clearly marked advising that the plates should not be exposed to X-rays. Also all plates in packets are equally affected and, unless an exceptionally high energy X-ray or γ -ray source was used for customs examination, the background of high energy electron tracks would not be obtained.

(d) *Radioactive Products from Atomic Explosions*

The last possibility is that the contamination is due to by-products of atomic explosions. There is apparently no significant overall increase in the activity in the air due to the recent series of explosions as most packets of plates received during the period of analysis show low electron background. However, there may be isolated patches or clouds of radioactive contamination which could account for the intermittent appearance of the heavy background.

In a personal communication Professor Vassy, University of Paris, has supplied details of the radioactivity found in rain or snow at Val-Joyeux, near Paris, this year, and results, which are to be published in *Comptes rendus de l'Académie des Sciences*, are shown in Table 2.

It will be observed from Table 1 that the most heavily exposed batch of plates left England on April 6 and the radiation which produced the exposure may have been due to the radioactive cloud which produced the radioactive rain in Paris on the same evening. It will also be seen that there is a peak in the rainfall activity from June 3 to 6 which links with the high electron background observed on the plate which left London on June 4. Plates which arrived in July and August have backgrounds in agreement with the lower activity of rainfall in these months.

Japanese scientists under Dr. Miyake (personal communication) also found that ash and sand which fell in Japan in mid April was highly radioactive and attribute radioactivity of the ash as due to an atomic explosion. They also observed peaks of activity in January and at the end of May which appear to link with the emulsion data.

Conclusions

Using the present method for importing plates by air to Australia there is a chance of 1 in 3 batches having too high a background. The reason for the high background has not been determined with certainty but it could have been produced by storing the plates too close to imported γ -ray isotopes. The films did not arrive with isotopes entering Australia and would need to be placed well within 4.5 ft of a box containing a 5.0 c ^{192}Ir source to obtain the highest backgrounds observed. Ruling out the effect due to isotopes there must be some other very highly active source encountered by the plates en route. The

date of departure from London of the most heavily affected packet agrees with the date of a highly radioactive downpour in France.

Whatever the cause of the anomalous backgrounds it is clear that before emulsions are used for experiments a test development is necessary to determine whether the plates are usable.

Beneficial discussions were had with Dr. C. E. Eddy, Mr. D. J. Stevens, and Mr. F. J. Richardson of the Commonwealth X-ray and Radium Laboratory.

AN UNUSUAL ELECTRON SHOWER*

By Y. K. LIM†

Figure 1 depicts an event observed in a 600 μ glass-backed Ilford G5 emulsion. The shower occurred at point 0 as indicated by a sudden increase in grain density of the incident track. For the first 150 μ from 0 the grain density of all secondary tracks (which are not separated) is 46 ± 4 grains per 50 μ , approximately 3.5 times the plateau value (13.0 grains/50 μ). 200 μ from 0 five tracks are resolved and the sum of their grain densities is 64 ± 2 grains/50 μ , approximately five times plateau. These values suggest that one track originated some 150–200 μ from 0, although the possibility that it originated at 0 is not entirely excluded.



Fig. 1.—Enlarged drawing of the electron shower.

Results of scattering measurements are given in Table 1, where the errors are statistical probable errors only. Tracks 5 and 6 can be identified as electrons, track 2 is probably an electron, and tracks 1, 3, and 4 are certainly not due to particles heavier than mesons. The most likely explanation is that all tracks are electrons as they are initially contained in a cone of half angle $\sim 0.5 \times 10^{-2}$ radians and this fact in itself suggests an electromagnetic interaction.

* Manuscript received August 29, 1955.

† Physics Department, University of Melbourne.

Assuming that all particles are electrons it can be seen that there is roughly an energy balance between the primary and secondary particles.

The event is consistent with the direct creation of an electron pair by a fast electron in the Coulomb field of an electron. Track 6 is then interpreted as a knock-on electron produced by one of these four electrons proceeding from 0, although the chance of this knock-on occurring so close to the origin is low and estimated as $1/500$. According to Heitler (1954) the expected ratio of pairs created in electron-electron collisions to pairs created in electron-nucleus collisions in a medium of atomic number Z is roughly $0.8/Z$. For nuclear

TABLE 1
MEASUREMENTS ON ELECTRONS OF SHOWER

Track	Length (mm)	$\bar{\alpha}$ (100) (deg/100 μ)	Energy (electrons) (MeV)
1	2.9	0.041 ± 0.011	650^{+230}_{-170}
2	2.2	0.20 ± 0.04	132^{+34}_{-21}
3	2.3	0.093 ± 0.023	285^{+94}_{-56}
4	2.3	0.13 ± 0.03	204^{+62}_{-38}
5	1.2	0.46 ± 0.11	57^{+19}_{-11}
6	1.4	2.0 ± 0.3	13 ± 2

emulsions this ratio is 0.023 . It is therefore curious that no event of this type has been previously reported, although a considerable number of tridents has been studied.

An analogous event interpreted as the direct creation of an electron pair in the field of an electron by a high energy incident photon has been reported (Hopper, Biswas, and Darby 1951), and again this is the only event of this type observed in the nuclear emulsions.

If we assume that all five tracks originated at 0 the result might be interpreted as double pair creation by a fast electron and the expected ratio of double pair creation to single pair creation is of the order of $1/137 \pi$ (Heitler 1954). Two events of the creation of a double pair by a photon in one elementary act have been reported previously (Hooper and King 1950).

The author wishes to thank Dr. V. D. Hopper for discussions on this event, which he had found.

References

- HEITLER, W. (1954).—"Quantum Theory of Radiation." 3rd Ed., pp. 228, 364. (Oxford Univ. Press.)
 HOOPER, J. E., and KING, D. T. (1950).—*Phil. Mag.* **41**: 1194.
 HOPPER, V. D., BISWAS, S., and DARBY, J. F. (1951).—*Phys. Rev.* **84**: 457.

THE ANGULAR SIZE OF THE VARIABLE RADIO SOURCE HYDRA-A*

By A. W. L. CARTER†

Slee (1955) has shown that the discrete cosmic radio source Hydra-A shows variations in intensity at metre wavelengths. This is the first variable radio source to be found, other than the Sun, although Ryle and Elsmore (1951) investigated about one hundred northerly sources and found no significant variations in their intensities over a period of 18 months. The observations to be described here were undertaken firstly to check the variability of the source, then suspected by Slee, and secondly to measure its angular size and see whether this might also be variable.

The measurements on any one day were made using either one or two radio interferometers chosen from three possible arrangements. It is well known that the brightness distribution across a radio source can be deduced from a number of interferometer observations at different angles and spacings (see, for example, Mills 1953). For the Hydra-A observations three aerials were available, each operating on a frequency of 101 Mc/s and spaced along an east-west line so that distances between pairs of about 90, 900, and 1000 wavelengths could be used.

The limited choice of spacings and angles used in the present observations does not allow the original brightness distribution to be reconstructed, so a simple model of the source has been assumed and the parameters of the model have been determined that best fit the observations. The model assumed has circular symmetry and a Gaussian distribution of radial intensity; this is a reasonable approximation to the central parts of some known radio sources and has the advantage that the derivation of an angular size and central brightness temperature is fairly simple.

The first observations were made in September 1954 while testing the equipment, when the apparent intensity of the source was measured with the 90- and 900-wavelength interferometers on alternate days. These measurements suggested a ratio between the intensities of 0.6 corresponding to a diameter of about $1\frac{1}{2}$ min of arc between half-brightness points. The next observations were made in early April 1955 as part of a survey of the angular sizes of the sources of medium intensity; they were also taken on different spacings on alternate days. Such observations, however, give unreliable angular sizes when made on a variable source, and so from the middle of May 1955 a systematic study of the Hydra-A source was begun, using one interferometer at a time for about 2 weeks continuously. Towards the end of June it became possible to use two interferometers at a time, and simultaneous observations with the 90-

* Manuscript received August 29, 1955.

† Division of Radiophysics, C.S.I.R.O., University Grounds, Sydney.

and 900-wavelength spacings were made until solar activity suspended this investigation in early July.

In all, 53 separate observations were made, many of which can be compared with simultaneous observations made by Slee. If the source is not resolved by the 90-wavelength interferometer, then the apparent intensity at this spacing should be directly comparable with Slee's intensities measured with the east-west arm of a Mills's cross aerial. Eleven such simultaneous observations showed positive correlation significant at the 5 per cent. level.

A similar comparison between the wide-spaced interferometer intensities and Slee's measurements can be used to test for variations in the angular size of the source. A group of 11 such observations with the 1000-wavelength interferometer showed negligible correlation. The fluctuations in intensities measured by each observer were much greater than the probable errors, so this lack of correlation suggests real variations in the apparent angular size. There is some suggestion that the major variations in the intensities were positively correlated.

More conclusive evidence for variations in the angular size is obtained from the final group of simultaneous observations with the 90- and 900-wavelength interferometers. These observations are considered the most reliable; the effect of ionospheric fluctuations was minimized by counting a large number of aerial lobes, and any changes in the properties of the aerial common to both interferometers do not influence the angular size measurements. The eight pairs of intensities measured by these two interferometers show positive correlation significant at the 10 per cent. level.

The intensities, angular sizes, and central brightness temperatures, derived from the last group of observations, are shown in Figure 1, and also the limits of error due to random effects of a single observation. The average angular size derived from these observations is 1.6 min of arc between half-brightness points and the average central brightness temperature is 6×10^5 °K. Any systematic error in the angular sizes is believed to be less than 10 per cent. The apparent negative correlation between the angular sizes and temperatures shown in the figure is a consequence of their functional relation and probably has little physical significance. The high value for the central brightness temperatures suggests that the radiation originates in a non-thermal process.

The temporal dependence and range of values of the variations in intensity and angular size derived from all the interferometer observations are of interest. The lowest total intensity measured is 30–40 per cent. of the mean value. The minima in total intensity are far more pronounced than the maxima, 20 per cent. above the average intensity being the largest value measured with the 90-wavelength interferometer. The minimum angular size measured is 1.0 min of arc and the maximum 2.0 min of arc, using Slee's measurements of the total intensities. There is no one-to-one correspondence between angular size and intensity but the major variations in these quantities appear to be positively correlated. No periodicity has been noticed in the variations of intensity or angular size.

On one occasion variations in total intensity were seen while the source was in the aerial beam. During this observation variations from about 75 to 35 per cent. of the average intensity were seen within 1 hr. No other observations on this source at the same time are available and no scintillations were seen. As ionospheric effects are a possible, but unlikely, cause of this variation, further observations will be needed before such rapid changes in intensity can be entirely accepted.

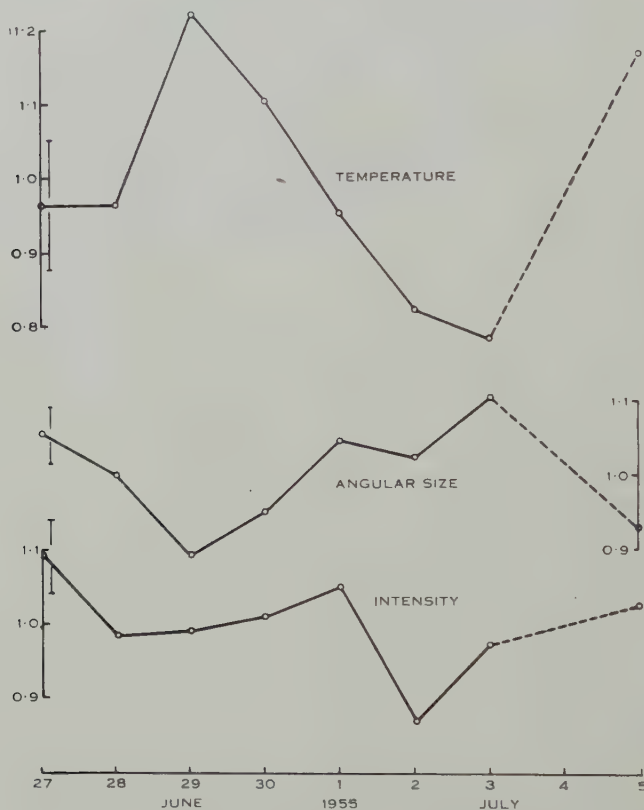


Fig. 1.—The intensity, apparent angular size, and central brightness temperature (normalized values) of the radio source Hydra-A from June 27 to July 5, 1955.

The variations of intensity and angular size may arise either in the source or in some intervening medium. Well-substantiated changes in intensity, from the mean value to a minimum, occur within a period of 1 day. This period, coupled with an overall apparent size of at least 1.6 min of arc, suggests that the variations, whatever the cause, must originate within about 2 parsecs from the solar system. It does not seem possible that some other brightness distribution, differing radically from the assumed model, could alter significantly this estimate of the maximum distance. For example, a distribution in which the variability is confined to a small, bright region of the source can be excluded, as such distributions would lead to a one-to-one correspondence between intensity

and angular size. Further evidence against such distributions is provided by the absence of changes in the apparent position of the source; no change in Right Ascension within the limits of detection, ± 2 min of arc, was detected. A small distant variable object, whose apparent size is increased by scattering in the intervening space, can also be excluded, as the apparent size of such an object would appear to remain constant. These considerations suggest that a considerable part of the apparent surface of the source must exhibit variations in brightness, and that the estimate of about 2 parsecs for the maximum distance from the Sun of the region in which the variations occur is unlikely to be seriously in error.

The present observations do not justify a more detailed analysis of possible mechanisms; they can be reconciled with quite simple obscuration effects in which the angular size of the original source of radiation is unlikely to be larger than the measured value. A consideration of intrinsic source variations is handicapped by the present lack of knowledge of the more general problem of radio-frequency generation in cosmic bodies. Some light might be thrown on possible mechanisms by systematic optical observations of this region and by measurements of the radio-frequency spectrum of the source.

The author is indebted to Mr. O. B. Slee for communicating his results prior to publication.

References

- MILLS, B. Y. (1953).—*Aust. J. Phys.* **6**: 452.
RYLE, M., and ELSMORE, B. (1951).—*Nature* **168**: 555.
SLEE, O. B. (1955).—*Aust. J. Phys.* **8**: 498.

INFORMATION ON ENERGY LEVELS IN ^8Be FROM NEUTRON ENERGY SPECTRA*

By J. R. BIRD† and R. H. SPEAR†

Introduction

A number of measurements have been made of the neutron energy spectra emitted in the reaction $^7\text{Li}(d,n)^8\text{Be}$, and, as is the case for other reactions leading to the same final nucleus, contradictory conclusions have been reached as to the number of levels observed.

Using nuclear emulsions as neutron detectors, early thick target experiments by Richards (1941) and Green and Gibson (1949) suggested the possibility of levels in ^8Be at 4.1, 5.3, and 7.5 MeV in addition to the well-known 3 MeV state. Trumpy, Grottdal, and Graue (1952), with somewhat better resolution,

* Manuscript received August 31, 1955.

† Physics Department, University of Melbourne.

obtained similar results, and suggested the existence of a further level at 2.2 MeV, while Catala, Aguilar, and Busquets (1953) claimed to have resolved levels at 1.5, 2.2, 2.9, 3.4, 4.1, 5.3, and 7.5 MeV. However, Trail and Johnson (1954*a*, 1954*b*), using a neutron spectrometer to obtain good statistics, found no evidence for levels below 10 MeV excitation other than the broad 3 MeV state. Another neutron spectrometer experiment by Reid (1954) produced evidence for the 5.3 MeV level; but Ihsan (1955*a*, 1955*b*), again using nuclear emulsions, reached similar conclusions to Trail and Johnson.

In view of the possibility of theoretical interpretation of the ^8Be level scheme a further study has been made of this reaction and the results of an investigation of the resolution achieved in the measurement of neutron energies have been used to analyse the results of this and previous experiments.

Experimental Results

A 50 keV thick target of separated ^7Li obtained from a small isotope separator was bombarded by 920 keV deuterons for an integrated target exposure of 25,000 μC . Neutrons were detected with 400 μ Ilford C2 emulsions, using the camera and target arrangement described by Bird (1955), and the plates

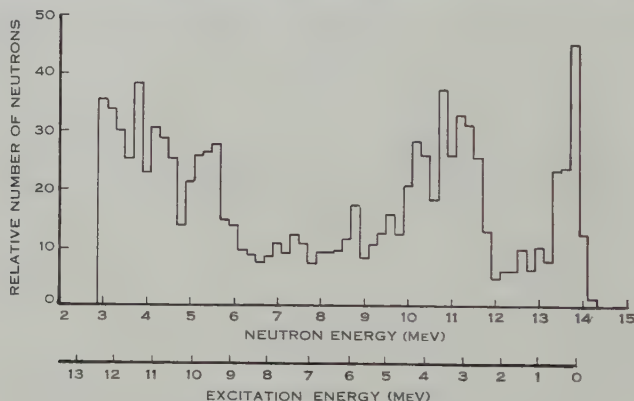


Fig. 1.—Energy spectrum of neutrons from $^7\text{Li}(d, n)^8\text{Be}$.

were studied using the same measuring procedure. The neutron energies were obtained using the range-energy relation given by Gibson, Prowse, and Rotblat (1954). The results for 1200 tracks measured at 90° are shown plotted in 200 keV intervals in Figure 1; and a further 400 tracks at 45° gave similar results. Corrections have been made for the loss of tracks from the emulsion and the variation of the neutron proton scattering cross section. The correction factor has been normalized to unity at 9.7 MeV and increases to 1.4 at 14 MeV.

The spectrum is similar to that published by Trumphy, Grottdal, and Graue (1952) and shows no indication of the fine structure reported by Catala, Aguilar, and Busquets (1953). In a comparison of the various experiments it is important to consider the resolution achieved in each case and to compare this with the performance necessary to distinguish between the neutron groups corresponding to the various proposed levels.

The Straggling of Protons and Neutrons in Nuclear Emulsions

The peak widths observed in the study of neutron energy spectra using nuclear emulsions contain contributions from the straggling of proton ranges and the uncertainties in measurement of track lengths and recoil angles.

The straggling of monoenergetic protons in Ilford C2 emulsions has been studied by various authors but the results obtained show considerable variations. The first measurements, summarized by Dyer (1952), suggest that the straggling expressed as a percentage of the total proton range is considerably higher than the corresponding values for air at low proton energies and approaches the air straggling at energies of 4–5 MeV. However, the results of Han and Endt (1954) are higher than the air straggling at 8–9 MeV, while those of Barkas, Smith, and Birnbaum (1955) at 33 MeV are lower.

Measurements have been made of the straggling of the proton tracks used by Bird and Hines (1954) in a study of the multiple scattering of monoenergetic protons. These tracks were obtained using a similar experimental procedure to that used by Dyer (1952), but with better geometry, giving a smaller spread in energy and direction of the protons. The range of approximately 300 tracks was measured at each of a number of proton energies, accepting only those tracks which were free from marked multiple scattering. The standard deviations of the observed range distributions were used to give the width at half height of the corresponding energy distributions.

The "Cellophane" absorbers, used to prevent scattered deuterons from reaching the photographic plate, contribute to the observed range straggling, and the results obtained using one, two, and three thicknesses of "Cellophane" were plotted against thickness and a smooth curve drawn through the points corresponding to each value of incident proton energy. The widths obtained by extrapolating each curve to zero thickness have been corrected for the effect of target thickness and are shown in Figure 2 (a), together with the results from other experiments and the equivalent air straggling (curve I, as given by Livingston and Bethe (1937)). The present results fall between those of Dyer and of Han and Endt, and are considerably higher than the air straggling. Curve II has been drawn to give the characteristic proton resolution for C2 emulsion, although its position is not uniquely defined, particularly at low energies.

Explanations of the greater straggling in emulsions than in air, have been given by Wilkins (1952) and Barkas, Smith, and Birnbaum (1955). Statistical fluctuations in the relative amounts of gelatin and silver halide traversed by each proton and the possible random movement of developed grains during the shrinkage of the processed emulsion will increase the straggling in each experiment by the same amount. However, the effects of processing techniques on the size and spacing of the developed grains will not be the same and may alter the observed straggling appreciably. For example, the grain spacing observed at the beginning of 5–10 MeV proton tracks, using the temperature method of development, would give an estimated increase in the straggling variance of about 1 per cent. of the mean range. This will affect the results of Han and Endt but not those of Dyer or the present straggling measurements, which were

obtained using a much stronger development. Also, the corrections for the effects of irregularities in absorber thickness, varying techniques in the selection and measurement of track lengths, and the straggling in energy of the incident particles have been made in different ways in each experiment, and may be responsible for part of the observed differences in results.

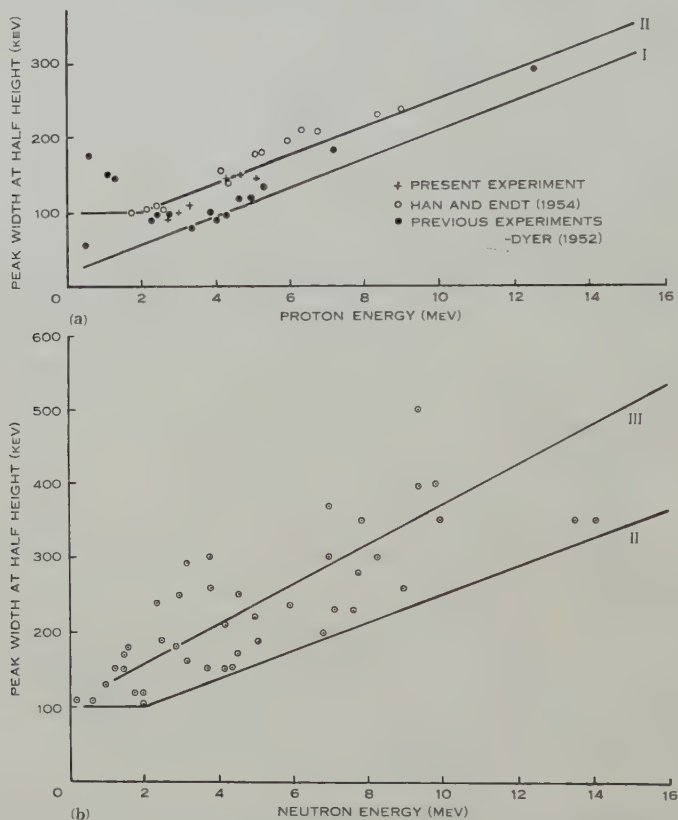


Fig. 2 (a).—Proton straggling. Curve I, equivalent air straggling; curve II, mean proton straggling.

Fig. 2 (b).—Neutron straggling. Experimental points from Gibson (1949), Nereson and Reines (1950), Ajzenberg (1952), Johnson (1952), Stelson and Preston (1952), Benenson (1953), Pruitt, Swartz, and Hanna (1953), Bird (1954), and Graue (1954). Curve II, as in Figure 2 (a); curve III represents average neutron straggling.

In the measurement of neutron energies by observation of recoil protons in nuclear emulsions an uncertainty of 2° in the measurement of each recoil angle will give an increase in straggling of the order of 10 per cent., but the actual increase will depend on individual measuring conditions. Figure 2 (b) shows values of the width at half height of neutron peaks from spectra published by a number of authors, plotted against neutron energy. These values have been corrected for the effects of target thickness. The results show large

variations at all energies and the average values represented by curve III are approximately 50 per cent. higher than the corresponding proton straggling (curve II from Figure 2 (a)). Although a number of the observed widths are close to the proton curve, many are so high that additional neutron straggling is required to explain them. The large widths of neutron peaks has been noted by Ajzenberg (1951) and is in agreement with the high values of the straggling variance required to explain the shape of thick target neutron peaks (Dyer and Bird 1953).

Any small angle scattering of the neutrons before they produce proton recoils will give an additional spread in energy and direction. The presence of scattering is shown by the asymmetry which is typical of many thin target neutron peaks, and this asymmetry has been found to increase with the distance of the area of measurement from the leading edge of the photographic plate, accompanied by a decrease in the estimated number of neutrons per unit solid angle. These changes are more marked at low energies owing to an increase in the neutron scattering cross section. The scattering which occurs in the photographic plate may often be more significant than the scattering in the material of the camera used for exposing the plates, and should be considered in the design of experiments requiring good resolution or the accurate measurement of neutron energies and angular distributions.

Discussion

Since the ground state of ^8Be is known to have a small natural width, the observed ground state peak width in each of the studies of $^7\text{Li}(d,n)^8\text{Be}$ may be used, in conjunction with Figure 2 (b), to indicate the resolution achieved in each case.

The use of thick targets by Richards (1941) and Green and Gibson (1949) is responsible for ground state peak widths of 600–800 keV and would prevent them from observing the fine structure suggested by Catala, Aguilar, and Busquets (1953). The latter experiment shows a ground state peak with a half width of approximately 300 keV, which falls below the proton straggling curve in Figure 2 (b), and this has not been achieved by any of the experiments used in the preparation of that figure. Since the experimental techniques used by these authors are similar to those used in other experiments, their suggestion of a fine level structure requires confirmation.

The ground state peak width observed by Trumpy, Grottdal, and Graue (1952) was 470 keV, and in the present experiment a width of 430 keV was obtained. These values fall close to curve III in Figure 2 (b) so that this curve may be expected to show the trend of peak widths expected at other neutron energies. This would be adequate to resolve the neutron groups suggested by Catala, Aguilar, and Busquets, and the lack of evidence for all these groups indicates that any excited states that occur in addition to the 3 MeV state must either have considerable widths or be present with small intensity relative to a continuous background of neutrons produced by alternative modes of breakup of the compound nucleus. The recent experiment of Ihsan (1955a, 1955b) gives a ground state peak width of 600 keV which would be barely sufficient to distinguish between proposed states with a separation of 700 keV.

The performance of neutron spectrometers cannot be analysed in this way, but the experiment by Trail and Johnson (1954*a*, 1954*b*) shows a ground state peak width of 1200 keV which would prevent the observation of fine structure in the neutron groups. Since the experiment by Reid (1954) covered only a small portion of the neutron spectrum it is difficult to compare his resolution with that of other experiments.

All the photographic plate studies of the neutron spectrum since that of Richards (1941) have shown irregularities suggesting the presence of excited states at 4.1, 5.3, and 7.5 MeV. The evidence from any one experiment is not conclusive, but the systematic appearance of these irregularities is significant.*

The authors are grateful to Professor L. H. Martin for his helpful interest in this work.

References

- AJZENBERG, F. (1951).—*Phys. Rev.* **82**: 43.
 AJZENBERG, F. (1952).—*Phys. Rev.* **88**: 298.
 BARKAS, W. H., SMITH, F. M., and BIRNBAUM, W. (1955).—*Phys. Rev.* **98**: 605.
 BENENSON, R. E. (1953).—*Phys. Rev.* **90**: 420.
 BIRD, J. R. (1954).—Ph.D. Thesis, University of Melbourne.
 BIRD, J. R. (1955).—*Aust. J. Phys.* **8**: 314.
 BIRD, J. R., and HINES, K. C. (1954).—*Aust. J. Phys.* **7**: 586.
 CATALA, J., AGUILAR, J., and BUSQUETS, F. (1953).—*An. Soc. Esp. Fis. Quim.* A **49**: 131.
 DYER, A. J. (1952).—*Aust. J. Sci. Res.* A **5**: 104.
 DYER, A. J., and BIRD, J. R. (1953).—*Aust. J. Phys.* **6**: 45.
 GIBSON, W. M. (1949).—*Proc. Phys. Soc. Lond.* A **62**: 586.
 GIBSON, W. M., and PROWSE, D. J. (1955).—*Phil. Mag.* **46**: 807.
 GIBSON, W. M., PROWSE, D. J., and ROTBLAT, J. (1954).—*Nature* **173**: 1180.
 GRAUE, A. (1954).—*Phil. Mag.* **45**: 1205.
 GREEN, L. L., and GIBSON, W. M. (1949).—*Proc. Phys. Soc. Lond.* A **62**: 407.
 HAN, K. K., and ENDT, P. M. (1954).—*Physica, 's Grav.* **20**: 311.
 IHSAN, M. A. (1955*a*).—*Phys. Rev.* **98**: 689.
 IHSAN, M. A. (1955*b*).—*Proc. Phys. Soc. Lond.* A **68**: 393.
 JOHNSON, V. R. (1952).—*Phys. Rev.* **86**: 302.
 LIVINGSTON, M. S., and BETHE, H. A. (1937).—*Rev. Mod. Phys.* **9**: 245.
 NERESON, N., and REINES, F. (1950).—*Rev. Sci. Instrum.* **21**: 534.
 PRUITT, J. S., SWARTZ, C. D., and HANNA, S. S. (1953).—*Phys. Rev.* **92**: 1456.
 REID, G. C. (1954).—*Proc. Phys. Soc. Lond.* A **67**: 466.
 RICHARDS, H. T. (1941).—*Phys. Rev.* **59**: 796.
 STELSON, P. H., and PRESTON, W. M. (1952).—*Phys. Rev.* **86**: 132.
 TRAIL, C. C., and JOHNSON, C. H. (1954*a*).—*Phys. Rev.* **95**: 1363.
 TRAIL, C. C., and JOHNSON, C. H. (1954*b*).—*Bull. Amer. Phys. Soc.* **29** (7): 34.
 TRUMPY, B., GROTDAL, T., and GRAUE, A. (1952).—*Nature* **170**: 1118.
 WILKINS, J. J. (1952).—A.E.R.E. Report G/R 1020.

* As this paper goes to press Gibson and Prowse (1955) have published preliminary measurements in which a resolution of 300 keV has been achieved. Levels are indicated at 2.1, 2.9, 4.05, and 5.25 MeV.

CORRIGENDUM

VOLUME 8, NUMBER 2, PAGES 248-72

"Theory of Tracks in Nuclear Research Emulsions"

By J. M. Blatt

Equations (3.20), (3.21), and (3.24) on p. 264 of this paper should be replaced by

$$\overline{n_r(t)} = pq^R(t/\beta - R - \frac{1}{2}), \dots\dots\dots (3.20)$$

$$\overline{x_r(t)} = pq^R \bar{w} t / [\beta + q^R [\bar{w}(1 - \frac{1}{2}p) - w_0(1 + Rp) - \beta q(R + 2p^{-1})], \dots\dots (3.21)$$

$$\mu_x^2 = \bar{w} + q[\beta^2/\bar{w}p^2 + 2q^R[\bar{w}(1 - \frac{1}{2}p) - w_0(1 + Rp) - \beta q(R + 2p^{-1})], \dots (3.24)$$

The changes in equations (3.20) and (3.21) are of no practical importance ; equation (3.24), however, is needed to estimate the error to be assigned to a measurement of $\overline{x_r(t)}$, and the change in the equation is significant. None of the conclusions of the paper are affected by this change, since this equation was not used in the discussion (Section IV).

The author would like to express his thanks to Drs. Castagnoli, Cortini, and Manfredini for calling these errors to his attention, and to them as well as to Mr. B. Chartres for checking the corrected expressions given above.

INDEX

	PAGE		PAGE
Aerial Smoothing, Correcting for Gaussian	54	Campbell, J. G.—	
Air Showers, Extensive, Observations on the Penetrating Component of	136	Photodisintegration of ^{16}O and ^{63}Cu by Photons of Variable Energy	449
Allen, T. K.— <i>See</i> Emeleus, K. G.	305	Carbon, Neutrons Emitted in the Deuteron Bombardment of	314
Aluminium Foils, Energy Loss Distributions of 1 MeV Electrons in	98	Caro, D. E., Martin, L. H., and Rouse, J. L.—	
Beryllium, Information on Energy Levels in ^8Be from Neutron Energy Spectra . .	567	A Variable Energy Cyclotron	306
Bird, J. R.—		Carter, A. W. L.—	
Neutrons Emitted in the Deuteron Bombardment of Carbon	314	The Angular Size of the Variable Radio Source Hydra-A	564
Bird, J. R., and Spear, R. H.—		Christiansen, W. N., and Warburton, J. A.—	
Information on Energy Levels in ^8Be from Neutron Energy Spectra	567	The Distribution of Radio Brightness over the Solar Disk at a Wavelength of 21 Centimetres. III. The Quiet Sun — Two - dimensional Observations . .	474
Blatt, J. M.—		Compression, Gravitational, of an Elastic Sphere	167
Theory of Tracks in Nuclear Research Emulsions	248	Computer, Proposed Modification to the C.S.I.R.O. Mark I	184
Boron, Excitation Studies of Reactions Occurring in the Proton Bombardment of ^{10}B	310	Computer, Starting Routines for the C.S.I.R.O. Mark I..	412
Bracewell, R. N.—		Copper, Photodisintegration of ^{63}Cu by Photons of Variable Energy	449
A Method of Correcting the Broadening of X-ray Line Profiles	61	Cornish, E. A.—	
Chord Construction for Correcting Aerial Smoothing	200	The Sampling Distributions of Statistics derived from the Multivariate t -Distribution	193
Correcting for Gaussian Aerial Smoothing	54	Corrigenda	318, 573
Correcting for Running Means by Successive Substitutions	329	Cosmic Rays and Air Mass Effects at Macquarie Island..	190
Bruce, C. F.—		Counter, Proportional, Design of a Multiple-wire, and its Use in the Study of the α -Particles from the Reaction $^7\text{Li} + p$	408
The Effects of Collimation and Oblique Incidence in Length Interferometers. I	224	Counter, Proportional, with Grid Control	419
Burr, E. J.—		Cowcher, E.— <i>See</i> Davies, L. W.	108
Sharpening of Observational Data in Two Dimensions	30		

	PAGE		PAGE
Csanady, G. T.—		Ellyett, C. D., and Roth, K. W.—	
Dispersal of Dust Particles		The Radar Determination of	
from Elevated Sources ..	545	Meteor Showers in the	
Current Layer, Hydromagnetic		Southern Hemisphere ..	390
Stability of a	319	Emeleus, K. G., and Allen,	
Cyclotron, A Variable Energy ..	306	T. K.—	
Das, S. C.—		Note on Plasma-Electron	
The Fitting of Truncated		Oscillations	305
Type III Curves to Daily		Equations, Linear Simultaneous,	
Rainfall Data	298	Solution by Matrix Iteration	425
Data, Observational, Sharpen-		Field Problems, Nets Composed	
ing of, in Two Dimensions ..	30	of Parts of Circles for the	
Davies, L. W., and Cowcher, E.—		Approximate Solution of ..	8
Microwave and Metre Wave		Fluids, Ionized, Eigen Oscilla-	
Radiation from the Positive		tions of Compressible ..	416
Column of a Gas Discharge	108	Fluorine, Photoprotons from ..	456
Davis, G.— <i>See</i> Herz, A. J. ..	129	Fraser, G. J.— <i>See</i> Ellyett, C. D.	273
Design, Statistical, of an Experi-		Galactic Survey at 400 Mc/s	
ment to Test the Stimulation		between Declinations -17°	
of Rain	440	and -49°	347
Deuterium, Diffusion of Slow		Galaxies, Observation and In-	
Electrons in	468	terpretation of Radio Emis-	
Deuteron Bombardment of		sion from Some Bright ..	368
Carbon, Neutrons Emitted in		Gas Discharge, Radio-frequency	
the	314	Radiation from the Positive	
Diffusion of Electrons in a Gas		Column of a	108
at Low Temperatures ..	551	Glass for Cemented Doublets,	
Diffusion of Slow Electrons in		A General Equation for the	
Deuterium	468	Choice of	68
Dislocations in Linear Arrays,		Guest, J.—	
Distribution of	1	The Solution of Linear Simul-	
Dust Particles, Dispersal from		taneous Equations by	
Elevated Sources	545	Matrix Iteration	425
Eccles, P. J.— <i>See</i> Officer, V. C.	136	Guest, P. G.—	
Elastic Sphere, Gravitational		A Differential Refractometer	
Compression of an	167	of High Sensitivity ..	212
Electron Trajectories, Deter-		Hall, Barbara I. H.—	
mination of, in the Presence		The Diffusion of Electrons in	
of Space Charge	74	a Gas at Low Temperatures	551
Electrons, Diffusion in a Gas at		The Diffusion of Slow	
Low Temperatures	551	Electrons in Deuterium ..	468
Electrons, Energy Loss Dis-		Hannan, E. J.—	
tributions of, Effect of		A Test for Singularities in	
Multiple Scattering on ..	90	Sydney Rainfall	289
Electrons, Energy Loss Dis-		Hanson, M. A.— <i>See</i> McDonell,	
tributions of 1 MeV, in		J. A.	98
Aluminium Foils	98	Harvey, J. A.—	
Electrons, Slow, Diffusion in		Movement of Sporadic <i>E</i>	
Deuterium	468	Ionization	523
Ellyett, C. D., and Fraser,			
G. J.—			
The Influence of Noise on			
Radar Meteor Observations	273		

	PAGE		PAGE
Head, A. K., and Louat, N.— The Distribution of Dislocations in Linear Arrays ..	1	Keane, A.— The Gravitational Compression of an Elastic Sphere..	167
Heat Transport through Narrow Channels by Liquid Helium II	206	Kerr, F. J., and de Vaucouleurs, G.— Rotation and other Motions of the Magellanic Clouds from Radio Observations	508
Hebbard, D. F., and Wilson, P. R.— The Effect of Multiple Scattering on Electron Energy Loss Distributions	90	Klemens, P. G.— The Transport of Heat through Narrow Channels by Liquid Helium II ..	206
Helium, High Temperature Excitation of	335	Laby, Jean E.— <i>See</i> Hopper, V. D.	557
Helium II, Liquid, Transport of Heat Through Narrow Channels by	206	Lasich, W. B., Muirhead, E. G., and Shute, G. G.— Photoprotons from Fluorine	456
Herz, A. J., and Davis, G.— The Characteristics of Tracks in Nuclear Research Emulsions	129	Lim, Y. K.— An Unusual Electron Shower	562
Hill, G. W., and Pearcey, T.— On Starting Routines for the C.S.I.R.O. Mark I Computer	412	Lithium, Use of a Multiple-wire Proportional Counter in the Study of the α -Particles from the Reaction ${}^7\text{Li} + p$..	408
Hollway, D. L.— The Determination of Electron Trajectories in the Presence of Space Charge..	74	Louat, N.— <i>See</i> Head, A. K. . .	1
Hopper, V. D., and Laby, Jean E.— Electron Background in Imported G5 Emulsions ..	557	Loughhead, R. E.— Eigen Oscillations of Compressible, Ionized Fluids..	416
Hydra-A, Angular Size of the Variable Radio Source ..	564	Hydromagnetic Stability of a Current Layer	319
Hydra-A, Apparent Intensity Variations of the Radio Source	498	McDonell, J. A., Hanson, M. A., and Wilson, P. R.— The Energy Loss Distributions of 1 MeV Electrons in Aluminium Foils ..	98
Interferometers, Length, Effects of Collimation and Oblique Incidence in. I	224	McGee, R. X., Slee, O. B., and Stanley, G. J.— Galactic Survey at 400 Mc/s between Declinations -17° and -49°	347
Interferometers, Length, Effects of Collimation and Oblique Incidence in. II	241	Magellanic Clouds, Rotation and other Motions of the, from Radio Observations ..	508
Ionosphere, Lunar Tide in Sporadic <i>E</i> at Brisbane ..	554	Martin, L. H.— <i>See</i> Caro, D. E.	306
Ionosphere, Movement of Sporadic <i>E</i> Ionization ..	523	Means, Running, Correcting for, by Successive Substitutions..	329
Jacklyn, R. M.— Cosmic Rays and Air Mass Effects at Macquarie Island	190	Meteor Observations, Influence of Noise on Radar	273
Jefferies, J. T.— The High Temperature Excitation of Helium ..	335	Meteor Showers in the Southern Hemisphere, Radar Determination of	390

	PAGE		PAGE
Meteor Trails, Diffusion Coefficients from the Rate of Decay of	279	Parthasarathy, R.— <i>See</i> Swarup, G.	487
Meteors in the Southern Hemisphere, Radio Echo Observations of	148	α -Particles from the Reaction ${}^7\text{Li}+p$	408
Microscopes, Reflecting, Multiple Images in	186	Pearcey, T.— <i>See</i> Hill, G. W. . . .	412
Mills, B. Y.— The Observation and Interpretation of Radio Emission from some Bright Galaxies	368	Photodisintegration of ${}^{16}\text{O}$ and ${}^{63}\text{Cu}$ by Photons of Variable Energy	449
Moran, P. A. P.— The Statistical Design of an Experiment to Test the Stimulation of Rain	440	Photoprotons from Fluorine	456
Muirhead, E. G.— <i>See</i> Lasich, W. B.	456	Positive Column of a Gas Discharge, Radio-frequency Radiation from the	108
Nets Composed of Parts of Circles for the Approximate Solution of Field Problems.. .. .	8	Proton Bombardment of ${}^{10}\text{B}$, Excitation Studies of Reactions Occurring in the	310
Neutron Energy Spectra, Information on Energy Levels in ${}^8\text{Be}$ from	567	Protons, Use of a Multiple-wire Proportional Counter in the Study of the α -Particles from the Reaction ${}^7\text{Li}+p$	408
Neutrons Emitted in the Deuteron Bombardment of Carbon	314	Quadratic Forms, Distribution of the Ratio of Two.. .. .	402
Norman, R. J.— A Proportional Counter with Grid Control	419	Radiation, Cosmic Radio, Angular Size of Hydra-A	564
Nuclear Emulsions, Characteristics of Tracks in	129	Radiation, Cosmic Radio, Apparent Intensity Variations of Hydra-A	498
Nuclear Emulsions, Electron Background in Imported G5	557	Radiation, Radio - frequency, from the Positive Column of a Gas Discharge	108
Nuclear Emulsions, Theory of Tracks in	248	Radio Emission from some Bright Galaxies, Observation and Interpretation of	368
Objectives, Cemented Doublet, A General Equation for the Choice of Glass for	68	Rain, Statistical Design of an Experiment to Test the Stimulation of	440
Officer, V. C., and Eccles, P. J.— Observations on the Penetrating Component of Extensive Air Showers	136	Rainfall Data, Fitting of Truncated Type III Curves to Daily	298
Oscillations, Eigen, of Compressible, Ionized Fluids	416	Rainfall, Test for Singularities in Sydney	289
Oscillations, Plasma-Electron, Note on	305	Refractometer, Differential, of High Sensitivity	212
Oxygen, Photodisintegration of ${}^{16}\text{O}$ by Photons of Variable Energy	449	Relaxation, Stress, of Wool Fibres in Water	176
		Rigby, B. J.— Stress Relaxation of Wool Fibres in Water at Strains of 5–20 per cent.	176

	PAGE		PAGE
Riviere, A. C., and Treacy, P. B.— The Design of a Multiple-wire Proportional Counter and its Use in the Study of the α -Particles from the Re- action ${}^7\text{Li}+p$	408	Taylor, R. J.— Some Observations of Wind Velocity Autocorrelations in the Lowest Layers of the Atmosphere	535
Roth, K. W.— <i>See</i> Ellyett, C. D.	390	t -Distribution, Multivariate, Sampling Distributions of Statistics derived from the..	193
Rouse, J. L.— <i>See</i> Caro, D. E...	306	Thomas, J. A., and Svenson, A. C.— Lunar Tide in Sporadic E at Brisbane	554
Sharpening of Observational Data in Two Dimensions ..	30	Thornton, B. S.— The Effects of Collimation and Oblique Incidence in Length Interferometers. II	241
Shower, Electron, An Unusual..	562	Tide, Lunar, in Sporadic E at Brisbane	554
Shute, G. G.— <i>See</i> Lasich, W. B.	456	Treacy, P. B.— <i>See</i> Riviere, A. C.	408
Slee, O. B.— Apparent Intensity Variations of the Radio Source Hydra-A	498	de Vaucouleurs, G.— <i>See</i> Kerr, F. J.	508
<i>See</i> McGee, R. X.	347	Warburton, J. A. — <i>See</i> Christiansen, W. N. ..	474
Smoothing, Aerial Chord Con- struction for Correcting ..	200	Watson, G. S.— The Distribution of the Ratio of Two Quadratic Forms ..	402
Spear, R. H.— <i>See</i> Bird, J. R...	567	Weiss, A. A.— Diffusion Coefficients from the Rate of Decay of Meteor Trails	279
Stability, Hydromagnetic, of a Current Layer	319	Radio Echo Observations of Meteors in the Southern Hemisphere	148
Stanley, G. J.— <i>See</i> McGee, R. X.	347	Wignall, J. W. G.— Excitation Studies of Re- actions Occurring in the Proton Bombardment of ${}^{10}\text{B}$	310
Steel, W. H.— A General Equation for the Choice of Glass for Cemented Doublets ..	68	Wilson, P. R.— <i>See</i> Hebbard, D. F. ..	90
Multiple Images in Reflecting Microscopes	186	<i>See</i> McDonell, J. A. ..	98
Sun, Distribution of Radio Brightness over the Solar Disk	474	Wind Velocity Autocorrelations in the Lowest Layers of the Atmosphere, Some Observa- tions of	535
Sun, Solar Brightness Distribu- tion at a Wavelength of 60 Centimetres	487	Wool Fibres, Stress Relaxation in Water	176
Svenson, A. C.— <i>See</i> Thomas, J. A.	554	X-ray Line Profiles, Method of Correcting the Broadening of	61
Swarup, G., and Parthasarathy, R.— Solar Brightness Distribution at a Wavelength of 60 Centimetres. I. The Quiet Sun	487		
Swire, B. E.— A Proposed Modification to the C.S.I.R.O. Mark I Computer	184		
Tasny-Tschiassny, L.— Nets Composed of Parts of Circles for the Approximate Solution of Field Problems	8		

DATE DUE

UIC AUG 18 '86

UIC REC'D AUG 18 '86

PERIODICALS MUST BE RETURNED
TO PERIODICALS DESK ONLY



3 8198 303 614 497

THE UNIVERSITY OF ILLINOIS AT CHICAGO

3 8198 303 614 497

